Vol. 69, No. 2 March 1962

Psychological Review

Richard L. Solomon, Editor
University of Pennsylvania

CONTENTS

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

(A large group of psychologists has kindly agreed to assist in the editing of manuscripts submitted to the *Psychological Review.* In each issue the Editor will list in the *Review* those editorial consultants who have helped him process manuscripts for that issue, so that their efforts may be gratefully acknowledged.)

The *Psychological Review* is devoted to articles of theoretical significance to any area of scientific endeavor in psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding 7,500 words will not be accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals. However, reports of original research may be included in a major theoretical article which attempts to integrate several related original studies.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts.* Detailed instructions for preparation of the abstract may be obtained from the Editor or from the APA Central Office.

All manuscripts must be submitted in duplicate.

Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors should retain a copy of the manuscript, in case there is loss in the mail. Manuscripts should be addressed to the Editor, Richard L. Solomon, Psychology Department, College Hall, University of Pennsylvania, Philadelphia 4, Pennsylvania.

Helen Orr
Acting Production Manager

Published Bimonthly by the
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Streets, Lancaster, Pa.
1333 Sixteenth St. N.W., Washington 6, D. C.

$10.00 volume $2.00 issue

Second-class postage paid at Lancaster, Pa.

Send all communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Vol. 69, No. 2 March 1962

PSYCHOLOGICAL REVIEW

TWO COMPONENTS IN BIPOLAR SCALES: DIRECTION AND EXTREMENESS[1]

DEAN PEABODY[2]

Columbia University

Various measures in psychology use a set of response categories extending in two directions about some neutral point. Scores for several responses are summed or averaged to give a composite score. Examples would include some of the methods of absolute judgment in psychophysics, the Likert-type of attitude scale, Osgood's semantic differential, rating scales used in personality assessment. The object of measurement for the composite scores may thus be some attribute of physical stimuli, of verbal concepts, of other persons, or of oneself.

Two components can be distinguished in the responses: (*a*) the direction from the neutral point, representing the basic dichotomy—heavy versus light, agreement versus disagreement with the statement, etc.; (*b*) the degree or extremeness of the response from the neutral point. For example, in a six-point Likert scale the subject writes + 1, + 2, + 3; or − 1, − 2, − 3 to indicate slight, moderate, and strong agreement or slight, moderate, and strong disagreement. The basic dichotomy of direction is determined by whether he agrees or disagrees (writes + or −); extremeness by whether he indicates slight, moderate, or strong feelings (writes 1, 2, or 3).

A neutral response category may or may not be provided—e.g., Likert scales traditionally permitted a neutral response (writing 0), but many recent Likert scales do not.

When scores for several responses are summed or averaged to give a composite score, the two components become: the number (more generally, the proportion) of responses scored in each direction; the mean extremeness of the responses in each direction.

The present paper derives a method for estimating the relative contribution of these two components to composite scores, and applies the method to data from several Likert attitude scales.

A possible reaction to the preceding might be to grant that the two components described can logically be distinguished, but question whether this is necessary or desirable. The purpose of using an extended scale is to in-

[1] This report represents a revision of a doctoral dissertation at Harvard University. The collection of data in England was aided by a grant from the Social Science Research Council; the replication in the United States by a grant from the Laboratory of Social Relations at Harvard. Support for rewriting was given by the Council for Social Sciences at Columbia. I am particularly grateful to G. W. Allport and F. Mosteller for general and statistical advice.

[2] Now at the University of Oregon.

crease the refinement of measurement over that given by dichotomous responses. What then is the purpose of separating direction and extremeness components? The answer is that there are both theoretical and practical reasons for considering the components separately:

1. The theoretical reason is that the distinction may be a psychological as well as a logical one. Ideally, the direction component would represent a qualitative aspect of the object of measurement, while the extremeness component would represent quantitative, intensity, aspects. In reality, both components are likely to represent, at least in part, response tendencies of the subjects rather than aspects of the object of measurement. Cronbach (1946; 1950, p. 21) distinguishes the response sets possibly involved in the two components as "bias" or "acquiescence" for the direction component, and "tendency to go to extremes" for the extremeness component. It seems desirable to examine these different factors separately. For example, Cronbach (1950) argues strongly that the extremeness component may largely represent a response set:

> The usual argument for the more finely divided scale of judgment on each attitude item is that it is more reliable and that subjects prefer it. If the latter advantage is significant, the finer scale may be retained and scored dichotomously. The argument that the finer scale gives more reliability is not a sound one, since this is precisely what we would expect if all of the added reliable variance were response-set variance and had no relation to beliefs about the attitude-object in question (p. 22).

2. The practical reason for examining the two components separately concerns the justification for dichotomous scoring. Regardless of the theoretical status of the two components, if the extremeness component is little represented in whatever the composite score measures, use of the simpler scoring method based on the direction dichotomy would seem justified.

Development of an Approximation for the Variance of Composite Scores in Terms of Direction and Extremeness Scores

The development given below assumes that a neutral response category is not provided for. Provision for a neutral category would give a closely analogous development.

The composite score used here represents the average of scores for the separate responses. An individual's composite score can be expressed mathematically in terms of direction and extremeness scores:

$$C = PX - (1 - P)\ Y \qquad [1]$$

where:
- C is the composite score.
- P is the proportion of responses in the direction scored positively.[3]
- X is the mean extremeness of responses in the direction scored positively.
- Y is the mean extremeness of responses in the direction scored negatively.

For example, a subject makes 10 responses as follows: $+2$, $+2$, $+2$, $+2$, $+3$, $+1$, -1, -1, -2, -3. If the responses are scored as they stand, then $P = .60$, $X = 2.00$, $Y = 1.75$, and $C = +.50$.

Frequently, to avoid negative numbers some constant is added to all scores. This simply raises the composite score and does not affect the development given below.

The present problem now becomes: to what extent does variation in the composite scores (C) depend on direction (P), and to what extent on extremeness (X and Y)?

[3] If a neutral response category were allowed, then Equation 1 should be modified to replace $1 - P$ by $Q =$ the proportion of responses in the direction scored negatively.

The variance of a difference in this case can be written as:

$$\sigma_C^2 = \sigma_{PX}^2 + \sigma^2_{(1-P)Y} - 2\ cov\ [PX,\ (1-P)Y] \quad [2]$$

The terms on the right involve the variances and covariance of products. There seems to be no simple method of handling this situation. The method used here is to approximate the terms on the right side of Equation 2 by the delta method.[4] In the present case, the assumption is that the deviations of individual scores around the group means are relatively small compared to the means, so that higher powers of the deviations can be neglected.

Let M_P, M_X, and M_Y stand for the group means of P, X, and Y. These means act as constants as regards variations between individual scores.

Let Δp, Δx, Δy stand for individual deviations around the group means:

$$P = M_P + \Delta p$$
$$X = M_X + \Delta x$$
$$Y = M_Y + \Delta y$$

These deviations are assumed to be small compared to the means.

The method will be illustrated for the first term on the right of Equation 2, σ_{PX}^2.

By the definition of any variance:

$$\sigma_{PX}^2 = E[(PX)^2] - [E(PX)]^2 \quad [3]$$

Applying the delta method:

$$PX = M_P M_X + \Delta p M_X + \Delta x M_P + \Delta p \Delta x \quad [4]$$

In taking expected values, the terms involving the first order of Δ vanish, since they are deviations around a mean, with an expected value of zero. The expected value of $\Delta p \Delta x$ is the covariance $(\Delta p,\ \Delta x)$. Hence:

$$E(PX) = M_P M_X + r_{(P,X)}\sigma_P \sigma_X \quad [5]$$

In squaring Equation 4 the terms are collected according to the order of Δ involved:

$(PX)^2 =$

0 order of Δ: $M_P^2 M_X^2$

1 order of Δ: $+ 2\Delta p M_P M_X^2 + 2\Delta x M_P^2 M_X$

2 order of Δ: $+ \Delta p^2 M_X^2 + \Delta x^2 M_P^2 + 4\Delta p \Delta x M_P M_X$

3 order of Δ: These terms will
4 order of Δ: be neglected.

In taking expected values, the first order terms again vanish. The third and fourth order terms are assumed to be negligible. The result is:

$$E[(PX)^2] = M_P^2 M_X^2 + \sigma_P^2 M_X^2 + \sigma_X^2 M_P^2 + 4r_{(P,X)}\sigma_P\sigma_X M_P M_X \quad [6]$$

Substituting from Equations 5 and 6 in Equation 3 gives:

$$\sigma_{PX}^2 = \sigma_P^2 M_X^2 + \sigma_X^2 M_P^2 + 2r_{(P,X)}\sigma_P\sigma_X M_P M_X - r^2_{(P,X)}\sigma_P^2\sigma_X^2 \quad [7]$$

Equation 7 gives an approximation for the first term on the right of Equation 2 in terms of the means, variances, and covariance of P and X.

For the second term on the right of Equation 2, an analogous application of the delta method gives the approximation:

$$\sigma^2_{(1-P)Y} = \sigma_P^2 M_Y^2 + \sigma_Y^2(1-M_P)^2 - 2r_{(P,Y)}\sigma_P\sigma_Y(1-M_P)M_Y - r^2_{(P,Y)}\sigma_P^2\sigma_Y^2 \quad [8]$$

For the third term on the right of Equation 2, the approximation is:

$$cov\ [PX,\ (1-P)Y] = -\sigma_P^2 M_X M_Y + r_{(X,Y)}\sigma_X\sigma_Y M_P(1-M_P) - r_{(P,X)}\sigma_P\sigma_X M_P M_Y + r_{(P,Y)}\sigma_P\sigma_Y M_X(1-M_P) + r_{(P,X)}r_{(P,Y)}\sigma_P^2\sigma_X\sigma_Y \quad [9]$$

[4] I am indebted to F. Mosteller for first working out the derivation leading from Equation 2 to Equation 10.

Substituting in Equation 2 from the approximations 7, 8, and 9 and collecting terms gives as the overall approximation:

$$
\begin{aligned}
&\qquad\quad \mathrm{I} \qquad\qquad\qquad\quad \mathrm{II_a}\\
\sigma_C^2 &= \sigma_P^2(M_X + M_Y)^2 + \sigma_X^2 M_P^2\\
&\qquad \mathrm{II_b}\\
&\quad + \sigma_Y^2(1 - M_P)^2\\
&\qquad \mathrm{II_c}\\
&\quad - 2r_{(X,Y)}\sigma_X\sigma_Y M_P(1 - M_P)\\
&\qquad \mathrm{III_a}\\
&\quad + 2r_{(P,X)}\sigma_P\sigma_X M_P(M_X + M_Y)\\
&\qquad \mathrm{III_b}\\
&\quad - 2r_{(P,Y)}\sigma_P\sigma_Y(1 - M_P)(M_X + M_Y)\\
&\qquad \mathrm{III_c}\\
&\quad - [r_{(P,X)}\sigma_P\sigma_X + r_{(P,Y)}\sigma_P\sigma_Y]^2 \qquad [10]
\end{aligned}
$$

Roman numerals are written above the terms to indicate the grouping that is relevant for present purposes: Term I involves variation of P; Terms $\mathrm{II_a}$, $\mathrm{II_b}$, and $\mathrm{II_c}$ involve variation of X and Y; Terms $\mathrm{III_a}$, $\mathrm{III_b}$, and $\mathrm{III_c}$ involve the covariation of P with X and Y.

(The final Term $\mathrm{III_c}$ involves a higher order of the variables than the other terms, and will often be negligible. The correlations, and the standard deviation of proportion scores are necessarily less than 1. The standard deviations of X and Y also tend to be less than 1 if the scale has no more than three steps in each direction —i.e., a scale of seven steps or less. In such cases the final Term $\mathrm{III_c}$ tends to be negligible. For scales that extend further, the term might not be negligible.)

The approximation given by Equation 10 can be used to estimate the relative contribution of direction and extremeness scores to the composite scores on different types of bipolar scales. Comparison can be made of the three groups of terms, and in particular, the relative size of Term I to the combined Term II.

Application to Data from Six-Step Likert Attitude Scales

This section applies the approximation just developed to data from six-step Likert attitude scales. For each item, the subject chooses between three degrees of agreement by marking + 1, + 2, + 3; or three degrees of disagreement by marking − 1, − 2, − 3. Hence X and Y must lie between 1 and 3, while the proportion scores P lie between 0 and 1.

Before taking up actual data, it is possible to get a general idea of the likely results. If one assumes for example that the means of P, X, and Y are near the middle of their possible values, then M_P will be around .5, and M_X and M_Y around 2. In this case:

Term I is about 16 σ_P^2.
Term $\mathrm{II_a}$ is about .25 σ_X^2.
Term $\mathrm{II_b}$ is about .25 σ_Y^2.
Term $\mathrm{II_c}$ is about $-.5\, r_{(X,Y)}\sigma_X\sigma_Y$.

Then with the standard deviations still to be evaluated, Term I has an advantage of 64 to 1 over Term $\mathrm{II_a}$ or Term $\mathrm{II_b}$. If one assumes further that the σ's are likely to be roughly in the ratio of their possible ranges, then σ_X and σ_Y would be about twice as large as σ_P. This cuts down the advantage of Term I over Term $\mathrm{II_a}$ or $\mathrm{II_b}$ to 16 to 1, or an advantage of 8 to 1 over both combined. Insofar as the correlation $r_{(X,Y)}$ is positive, the combined Term II would be made even smaller by Term $\mathrm{II_c}$.

While in principle, data from any set of items could be used, two points deserve consideration.

1. If the items were such that subjects tended to respond in the same direction on nearly all items, then

TABLE 1

BASIC DATA FOR DIRECTION AND EXTREMENESS SCORES

Score	Scale							
	F		Dogmatism		(balanced) Conservatism		General Agreement	
	US	UK	US	UK	US	UK	US	UK
Direction: P								
M_P	.45	.50	.45	.50	.59	.45	.58	.62
σ_P	.12	.18	.10	.13	.11	.16	.05	.06
Rel. of P	.56	.78	.51	.67	.50	.84	.75	.89
Extremeness: X & Y								
M_X	1.79	1.81	1.74	1.75	2.15	1.97	2.01	1.96
σ_X	.34	.31	.29	.28	.35	.36	.25	.23
Rel. of X	.66	.57	.62	.53	.80	.71	.94	.88
M_Y	1.96	1.90	1.94	1.83	1.80	1.95	1.95	1.84
σ_Y	.33	.36	.32	.33	.33	.33	.29	.27
Rel. of Y	.64	.70	.71	.67	.66	.64	.92	.91
Correlations								
$r_{(P,X)}$	.08	.48	.11	.22	.40	.11	−.07	.00
$r_{(P,Y)}$	−.20	−.24	−.08	.01	−.12	−.38	−.14	−.16
$r_{(X,Y)}$	.45	.12	.46	.47	.44	.47	.89	.82

either X or Y would be based on few items and would tend to be an unstable measure.[5]

2. To give a representative test, a variety of types of scales should be used. In particular, the test should include both the recent type of scale deriving from *The Authoritarian Personality* where all items are worded in one direction and agreement is always scored positively, and the more traditional type of Likert scale where items are worded in both directions and agreement is sometimes scored positively and sometimes negatively.

Accordingly results are presented for four sets of items:

1. Responses to 28 items from the F Scale of "authoritarianism" (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950, pp. 255–257, 260—omitting Items 22 and 30/28

2. Responses to the 40 items of Rokeach's Dogmatism Scale, Form E (Rokeach, 1956, pp. 7–10)

In both of these scales all items are worded in one direction and agreement is always scored positively. Hence scores represent a mixture of attitudes toward Authoritarianism or Dogmatism with an agreement response tendency ("acquiescent response set").

3. Responses to 32 items referring to conservative versus more leftist positions on economic and political issues. Sixteen of the statements present a conservative position and agreement is scored positively. The remaining 16 statements present the same issue in the reverse direction (the two versions are given at different times) and agreement is scored negatively. Hence the combined score should cancel major effects of acquiescent response set and give a fairly pure measure of attitudes toward Conservatism.

4. Responses to 208 items, scored for "General Agreement." The items include originals and reversals of the F and Dog-

[5] In the limiting case, if a subject responds to all items in one direction, either his X or his Y score becomes indeterminate.

TABLE 2

CONTRIBUTION TO VARIANCE OF COMPOSITE SCORES BY DIRECTION AND EXTREMENESS SCORES

Terms from Equation 10	Scale							
	F		Dogmatism		(balanced) Conservatism		General Agreement	
	US	UK	US	UK	US	UK	US	UK
I	.193	.439	.138	.213	.195	.398	.039	.056
II_a	.024	.025	.017	.019	.041	.027	.021	.020
II_b	.032	.032	.031	.026	.018	.033	.015	.011
II_c	−.025	−.007	−.021	−.021	−.024	−.028	−.032	−.024
III_a	.011	.100	.011	.028	.072	.023	−.004	.000
III_b	.031	.057	.010	−.001	.015	.085	.007	.008
II combined	.031	.050	.027	.024	.035	.032	.004	.007
III combined	.042	.157	.021	.027	.086	.108	.003	.008
Total: I + II + III	.265	.646	.186	.265	.316	.538	.046	.071
Actual σ_C^2	.276	.642	.195	.268	.300	.585	.048	.066
% discrepancy	4%	0.7%	5%	1%	5%	8%	2%	7%
% of Total:								
I: Variance of P	73	68	75	81	62	74	85	79
II: Variance of X & Y	12	8	14	9	11	6	10	10
III: Covariance of P with X & Y	16	24	11	10	27	20	6	11

matism items, as well as for 20 items from the Anti-Semitism scale (Adorno et al., 1950, pp. 68–70); and both versions of the Conservatism items just mentioned. Agreement is scored positively for all items. Hence the combined score should cancel major effects of attitude content and give a measure of generalized "acquiescent response set."

The subjects were 88 American and 75 English engineering college students. Results for the American sample are headed "US"; results for the English sample "UK." Further details of procedure are given elsewhere (Peabody, 1961).

Table 1 gives the means, standard deviations, reliabilities,[6] and correlations for the direction and extremeness scores P, X, and Y.

[6] All reliabilities are for split halves separated by a temporal interval, usually 2 weeks.

Table 2 gives the evaluation of the terms of the approximation Equation 10. (The final Term III_c is omitted here since of negligible size.)

The closeness of the approximation is given in Table 2 by the "% discrepancy" between the sum of the terms and the actual variance of the composite scores, and ranges from around 1% to 8%. This result is about what would be expected from this type of approximation.

Although the results in Table 2 differ in detail from the rough a priori estimate made of the relative importance of Terms I and II, the final conclusion is about the same. The average result for the two samples on each set of items gives the following composition of the variance of the com-

posite scores:

About 70%–80% involves variation of direction scores—Term I
About 10% involves variation of extremeness scores—Term II
About 10%–20% involves covariation of direction scores with extremeness scores—Term III

This is a ratio of 7–8 to 1 for the first two terms.

Two main factors seem responsible for this result:

1. From the general form of Equation 10 variation in one component is weighted by the second power of the mean(s) from the other component (i.e., whether a person agrees or disagrees with an item will usually have a larger effect on his composite score than the degree of his agreement or disagreement). As in the a priori estimate, these weights overwhelmingly favor the direction variance term and more than compensate for the smaller size of this variance itself—so that Term I is considerably larger than any of the Terms II.

2. The correlation between extremeness scores in the positive and negative directions is always positive. Those who use large positive numbers (strongly agree) tend to use large negative numbers (strongly disagree). Hence Term II_c cancels much of the effect of the extremeness component. This factor becomes particularly prominent for the General Agreement scores where the variance of direction scores is relatively small (presumably due to the balancing out of content), but the correlation between X and Y becomes very substantial.

A simpler if less rigorous approach is to consider the more practical question: how good a substitute for whatever is measured by the composite scores can be given by direction or by extremeness scores? Table 3 gives the relevant correlations together with the reliabilities of the composite scores.

The correlations given in Table 3 are representative of the values found for a wider variety of scales. The correlations of composite scores with direction scores tend to be around .90 and above. The correlations with the extremeness scores are more variable and substantially lower.

The results show fairly clearly that composite scores on these Likert attitude items reflect primarily the direction of responses, and only to a minor extent their extremeness. The practical conclusion suggested is that there

TABLE 3
Data for Composite Scores

Reliabilities and correlations	Scale							
	F		Dogmatism		(balanced) Conservatism		General Agreement	
	US	UK	US	UK	US	UK	US	UK
Reliability of C scores	.69	.82	.62	.68	.66	.86	.77	.66
Correlations of C scores:								
with Direction scores—								
$r_{(C,P)}$	.93	.97	.93	.96	.87	.97	.96	.95
with Extremeness scores—								
$r_{(C,X)}$	.15	.56	.16	.31	.52	.16	.03	.23
$r_{(C,Y)}$	−.41	−.38	−.38	−.15	−.19	−.47	−.18	−.06

is justification for scoring such items dichotomously according to the direction of response. The resulting score will closely reflect whatever the composite score would.

Further applications are necessary to determine whether these conclusions would extend to scales where a neutral category is permitted, to scales with more than six or seven steps, and to scales that do not concern attitudes. As regards more extended scales (e.g., of 9 or 11 steps) the reasoning used with regard to six-point scales would imply that the extremeness component might play a more significant role. However, in view of the prevalence in psychology of scales with no more than six or seven steps, the present results may be widely applicable.

Discussion

Extremeness scores are quite reliable if based on a large enough number of responses, and are to a considerable extent independent of direction scores. But although extremeness scores can be stable measures of something, they are largely unrepresented in the usual composite scores. Hence the hope that Likert scale scores for example should be "influenced by the *degree* to which subjects favor or oppose attitude statements" (Newcomb, 1950, p. 171) is largely unfulfilled in practice.

Extremeness scores will tend not to have the same correlates as the composite scores, but could have correlates of their own. Only a few studies using separate measures of extremeness give some evidence of what these correlates may be (e.g., Osgood, Suci, & Tannenbaum, 1957, pp. 226–236; Soueif, 1958). In particular, the general conclusion that extremeness is little reflected in whatever the composite scores measure applies to the F Scale of "authoritarianism." This fact is worth noting in view of the recurrent attempts to interpret extremeness as "intolerance of ambiguity," implying that extremeness ought to relate to composite scores on the F Scale—most recently Mogar (1960) uses extremeness on the semantic differential as a criterion for composite F Scale scores.

What is the nature of extremeness scores themselves? Two possibilities have already been suggested: extremeness scores may represent actual differences in intensity (e.g., of attitudes), or they may only represent differences in using the response scale (i.e., response set). Differences within an individual in the extremeness of responses to different items presumably represent primarily differences in intensity, although even here the individual may change his usage of the response scale over time.

On the other hand, as regards differences between individuals in average extremeness, a different interpretation is suggested by the correlations between extremeness scores found in the study reported here. Extremeness seems to be a very general individual characteristic: correlations between extremeness scores in both directions for all possible sets of items (including originals and reversals of the F, Dogmatism, Anti-Semitism, and Conservatism items) tend all to be positive, most of them substantially so. Thus, a person who shows relative extremeness in taking an "authoritarian" position when agreeing with F Scale originals or disagreeing with reversals, tends also to show relative extremeness in taking an "anti-authoritarian" position when disagreeing with F Scale originals or agreeing with reversals, etc. It does not seem very plausible to assume that individual differences in the intensity of attitudes show such systematic generality for responses in opposite directions. But

the size of the numbers subjects actually write does show such generality. Hence differences in extremeness between persons seem to represent primarily response sets characteristic of the individual—consistent tendencies to use extreme or moderate response categories. While some evidence suggests a secondary tendency for extremeness scores to represent differences in intensity,[7] the primary factor —the very great generality of extremeness scores—seems better interpreted as response set.

SUMMARY

Scores on the bipolar scales in frequent use in psychology can be analyzed into two components: the proportion of responses in each direction of the basic dichotomy, and the extremeness of responses in each direction. An approximation is developed expressing variation in the usual composite scores in terms of variation in the two components of direction and extremeness.

The approximation is applied to four sets of items from Likert attitude scales. The results indicate that composite scores reflect primarily the direction of responses, and only to a minor extent their extremeness. The practical implication is that there is justification for scoring the items dichotomously according to the direction of response.

Since extremeness scores are reliable and are largely unreflected in the usual composite scores, they may have quite different correlates of their own. Individual differences in average extremeness show wide generality across different extremeness scores. This generality extends to responses in opposing directions—suggesting that the differences primarily represent response sets, and only to a secondary degree actual differences in intensity.

[7] For example, correlations tend to be positive between P and X, and negative between P and Y (as in Table 1). This indicates some tendency to use higher or lower positions on the response scale as a whole, as would be expected if intensity were a factor. However, this tendency should lead to negative correlations between X and Y, and is in every case overcome by the opposing tendency for extremeness scores to correlate positively.

REFERENCES

ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality.* New York: Harper, 1950.

CRONBACH, L. J. Response sets and test validity. *Educ. psychol. Measmt.*, 1946, **6**, 475–494.

CRONBACH, L. J. Further evidence on response sets and test design. *Educ. psychol. Measmt.*, 1950, **10**, 3–31.

MOGAR, R. E. Three versions of the F scale and performance on the semantic differential. *J. abnorm. soc. Psychol.*, 1960, **60**, 262–265.

NEWCOMB, T. M. *Social psychology.* New York: Dryden, 1950.

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning.* Urbana: Univer. Illinois Press, 1957.

PEABODY, D. Attitude content and agreement set in scales of authoritarianism, dogmatism, anti-semitism, and economic conservatism. *J. abnorm. soc. Psychol.*, 1961, **63**, 1–11.

ROKEACH, M. Political and religious dogmatism: An alternative to the authoritarian personality. *Psychol. Monogr.*, 1956, **70**(18, Whole No. 425).

SOUEIF, M. I. Extreme response sets as a measure of intolerance of ambiguity. *Brit. J. Psychol.*, 1958, **49**, 329–334.

(Received August 23, 1960)

Psychological Review
1962, Vol. 69, No. 2, 74–90

THE LATERAL HYPOTHALAMIC SYNDROME:

RECOVERY OF FEEDING AND DRINKING AFTER LATERAL HYPOTHALAMIC LESIONS [1]

PHILIP TEITELBAUM AND ALAN N. EPSTEIN [2]

University of Pennsylvania

The brain controls feeding behavior. If the base of the brain is damaged bilaterally in the ventromedial region of the hypothalamus, rats and other mammals overeat and become tremendously obese (Anand, Dua, & Schoenberg, 1955; Brobeck, Tepperman, & Long, 1943; Wheatley, 1944). Damage in the lateral hypothalamus produces the opposite effect: the animals do not eat and will die in the presence of food and water (Anand & Brobeck, 1951; Anand, Dua, & Schoenberg, 1955). Electrical stimulation through chronically implanted electrodes in the medial hypothalamus suppresses feeding, and stimulation in the lateral hypothalamus elicits it (Brügger, 1943; Larsson, 1954; Wyrwicka & Dobrzecka, 1960).

These dramatic and consistent phenomena suggest that a dual hypothalamic mechanism controls food intake. Cessation of eating (satiety) is said to be controlled by a medial neural system whereas active feeding is maintained by a system in the lateral hypothalamus (Anand & Brobeck, 1951).

There have been two major developments since the original description of the lateral starvation phenomenon by Anand and Brobeck in 1951 which suggest a reinterpretation of the functions of the lateral hypothalamic system. First, not only do the animals refuse to eat, but they also do not drink (Teitelbaum & Stellar, 1954). The adipsia was emphasized by Montemurro and Stevenson (1957). They reported three rats with lateral hypothalamic lesions that did not drink water but did eat appreciable quantities of a high-fat diet, maintaining their weight for as long as 55 days when tube-fed with 10% glucose solution or water. Without hydration, the animals ate less and lost weight. This was the first demonstration that lateral hypothalamic lesions in the rat could produce adipsia without aphagia. The more recent finding that hypertonic saline injected directly into the lateral hypothalamus can elicit drinking as well as eating (Epstein, 1960a) is consistent with this.

Second, the refusal to eat and drink produced by these lesions is not always permanent, and when recovery occurs, it follows a regular and orderly course. Teitelbaum and Stellar (1954), working with large lateral hypothalamic lesions that produced both aphagia and adipsia, kept their rats alive by tube-feeding. Without maintenance

[1] This work was supported by NSF Grant G-9792 and by USPHS Grant 2 b-5273. A preliminary report was presented at the 127th meeting of the AAAS in New York on December 29, 1960.

The authors are grateful to Eliot Stellar for his generous advice and material support throughout this work. Both he and R. Duncan Luce made valuable criticisms of the manuscript. Malcolm M. Cohen and Jay J. Singer gave valuable technical assistance.

[2] Formerly at the Institute of Neurological Sciences, Medical School, University of Pennsylvania. Now at the Zoological Laboratories, Division of Biology at the university.

these animals would have starved to death, but when kept alive they eventually began to eat and drink again. The animals first accepted wet and palatable foods and began to maintain their weight, then drank water, and finally, regulated their weight on ordinary dry food and water, and seemed normal.

We will show that lateral hypothalamic lesions produce a complex set of symptoms—a syndrome of deficits in drinking as well as in feeding behavior from which the animals recover at different rates. In 35 rats (34 females, 1 male), we have produced bilateral lateral hypothalamic lesions of varying sizes using a standard stereotaxic technique (Stellar & Krause, 1954). The coordinates used were: A 6.0, RL 2.0, and 8.0 down from the dural surface. By offering the animals a variety of wet and palatable foods[3] in addition to the ordinary dry Purina laboratory chow pellets and water, and by keeping them alive by tube-feeding if necessary, we have identified four clear-cut stages in the recovery from lateral hypothalamic lesions. The stages always follow the same sequence and are defined by critical behavioral events that occur in succession during the recovery process. Each of the stages has been analyzed experimentally in an attempt to identify the physiological deficits underlying it.

[3] A liquid diet was prepared by mixing 75 milliliters of whole egg; 125 milliliters of 50% sucrose solution; 240 milliliters of evaporated milk; 0.3 milliliters Poly-Vi-Sol (multiple vitamin preparation of Mead-Johnson Co., Evansville, Indiana); and 30 milliliters Kaopectate (Upjohn Co., Kalamazoo, Michigan). One part 10% formalin in 100 milliliters of diet was added as a preservative. Milk chocolate refers to finely divided Hershey's milk chocolate. A wet-mash was prepared by mixing Purina Laboratory Chow Powder in equal parts with water.

The animal whose record is graphed in Figure 1 had large symmetrical lesions which resulted in a long, slow recovery, thus making it possible to see the events in each stage quite clearly.

STAGES IN RECOVERY

Stage I: Aphagia and Adipsia

The animals refuse all food, wet or dry, and they do not drink any water. Without tube-feeding, they will die. The animal of Figure 1 neither ate nor drank for 19 days, steadily lost weight, and had to be kept alive from the 7th day on by gastric intubation with a liquid diet. The behavior of such a rat is quite characteristic. It is generally inactive, but it is alert and can be easily roused to movement. It walks with a slightly hunched posture, carrying its head low to the ground. It typically ignores food and will actively push away food that is forced upon it. If such a rat gets wet food on its paws or face, it engages in vigorous bilateral shaking of the front paws which somewhat resembles the paw- and face-grooming movements of the normal rat except that there is no grooming with the mouth. When its paws are wet, it wipes them repeatedly on the floor of the cage, and if the face and mouth have food or water on them, it walks along and rubs its chin and the sides of its face on the cage floor. It does not groom itself effectively, allowing its fur to become completely matted and caked with wet food. It actively resists having milk placed in its mouth by a medicine dropper, and it does not swallow the milk once it is there, but rather allows it to dribble out the side of the mouth. Ordinarily a normal rat does not show such behavior. But it does engage in the same paw-waving and wiping, chin-rubbing, poor grooming,

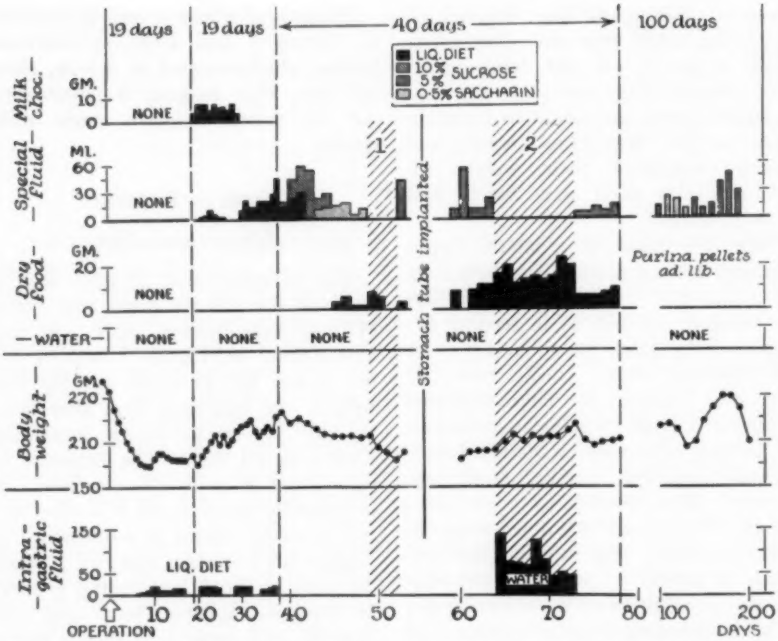


FIG. 1. Recovery of food intake with permanent adipsia after large lateral hypothalamic lesions. (The first two dashed vertical lines set off stages of recovery. First, for 19 days, the animal neither eats nor drinks; second, for 19 more days, it eats wet and palatable foods but still has to be tube-fed; and third, for the rest of its life, it regulates its food intake but does not drink water. The rat was weaned to sweet nonnutritive fluid and then ate dry food. When offered only water, Shaded Area 1, it did not drink and then stopped eating. Gastric hydration, Shaded Area 2, allowed normal feeding. Thereafter, the animal ingested sweet fluid and survived on dry food.)

and rejection when very bitter quinine (1% weight/volume) is put in its mouth or on its paws, face, and fur. This suggests that mouth contact with food and water is highly aversive to a rat with lateral lesions during this stage.

As such a rat loses weight each day, its posture becomes increasingly more hunched, the fur is piloerected, and its activity may increase. This is similar to the effect of starvation in a normal animal and may result from the difficulty a rat has in maintaining its body temperature without food (Stevenson & Rixon, 1957). The feces of a rat with lateral hypothalamic lesions are often mucoid and may be quite loose and tarry. When tarry stools are present, a test for occult blood in the feces is usually positive.[4] If tube feeding is withheld and the rat dies, after 7 to 10 days, we have often found massive ulcers in the stomach and hemorrhages in both stomach and cecum (Andersson & McCann, 1956; French, Porter, von Amerongen, & Raney, 1952). However, this does not seem to be the reason they refuse to eat and drink, because in some aphagic and

[4] Performed with Hematest (Ames Co., Elkhart, Indiana) tablets.

adipsic animals there has been no macroscopic evidence of gastrointestinal pathology.

Stage II: Anorexia and Adipsia

The animals eat wet and palatable foods, such as a liquid diet, milk chocolate, or cookies,[5] but refuse dry pellets and do not drink water. Although they eat appreciable quantities of wet and palatable foods, they do not eat enough to maintain their body weight, and they die unless tube-feeding is continued. The animal is therefore anorexic. It eats but it does not eat enough to keep itself alive.

The rat in Figure 1 nibbled at milk chocolate and liquid diet beginning on the 20th day, but for an additional 19 days it did not eat enough to maintain its weight without tube-feeding. Along with these changes in feeding go many other changes in the animal's behavior. It begins to use its mouth in grooming itself and its fur gradually is cleaned of matted hair and dried food. Now when fresh food is placed in the cage, the rat no longer ignores it completely. The animal actively approaches and investigates odorous and palatable foods such as milk chocolate and the liquid diet. If the food is liquid, it may dip its paws in it, or even take an occasional lick or bite at it, particularly around the rim of the container. If it is a wet mash, the rat may paw it and scatter it around on the floor or nibble at it a few times. It often chews vigorously as it stands over the food or walks around the cage, yet little or no food is eaten. Palatable and odorous foods are most effective in stimulating the animal to investigate and eat a little, but the eating is not sustained. It is as though the odor, taste, and other qualities of the diet arouse the animal's appetite, but there is no maintained character to the eating. When a hungry normal rat is presented with fresh palatable food, it gulps it down vigorously, eating a great quantity in a few minutes. A lateral hypothalamic animal in this stage often approaches the food, pauses over it, and then eats in a rather delicate and gingerly fashion even though it is quite starved. The vigorous urge to eat seen in the normal animal is still lacking in the animal with lateral hypothalamic lesions.

But hunger gradually returns. This may be seen most clearly in an animal which exhibits prolonged anorexia and is kept alive by tube-feeding. Each day its interest in fresh food grows more active and its chewing and licking more prolonged. In the intervals between tube-feedings, its activity begins to increase so that the animal sometimes displays the hyperactivity characteristic of a starved but otherwise normal rat. Sometimes associated with this is a pronounced hypothermia (as low as 96 degree Fahrenheit instead of the normal 99 to 100 degree Fahrenheit) if the body weight is very low. The activity diminishes if the animal is warmed by a lamp directly over its cage. As the animal nibbles more, it needs less tube-feeding to maintain its weight. Then, one day, it eats a large quantity of food (39th day after operation for the animal shown in Figure 1), and its weight takes a sudden upward spurt. It is now no longer anorexic.

Stage III: Adipsia with a Secondary Dehydration-Aphagia

The rats now regulate their caloric intake of wet and palatable foods, but they still refuse to drink water (Williams & Teitelbaum, 1959). They ingest a liquid diet vigorously, and

[5] Sunshine Chocolate Chip cookies (Sunshine Biscuits, Inc., Long Island City, New York).

their coats grow sleek as their weight rises, often to preoperative levels. Palatability is no longer essential for food intake. If these rats are trained to press a bar that allows a liquid diet to flow directly into the stomach through a chronically implanted gastric tube (Epstein, 1960b), they ingest normal amounts of food and maintain their weight throughout a 6-day test. They regulate their caloric intake normally even when the diet is diluted with water. When required to press the bar three times for each injection of food, they work sufficiently harder to maintain the intake. This regulation of food intake without stimulation of the oropharyngeal receptors has been demonstrated in three rats during this stage and is similar in all essential respects to that seen in normal animals feeding themselves intragastrically (Epstein & Teitelbaum, in press). The animals press the bar because they are hungry, and the taste of the food is no longer necessary to elicit feeding.

The animals' feeding behavior seems normal, but they do not drink water. However, they can be weaned from the liquid diet to a sweet nonnutritive solution. By mixing the liquid diet each day with an increasing proportion of the sweet fluid, they can be weaned first to a 10% weight by volume (w/v) sucrose solution, and from this to a 0.5 or 0.2% w/v sodium saccharin solution. Thus, for the first time since the lateral hypothalamus was damaged, they are freely drinking nonnutritive fluid while still refusing dry food and water. As their weight begins to drop, they show the hyperactivity of starvation, then they inevitably begin to eat the dry Purina diet and regain their weight. If the saccharin solution is then removed (as indicated in the hatched area labeled 1 starting on the 49th postoperative day in Figure 1), they still refuse to drink water, their food intake drops to zero in 2 or 3 days, they lose weight, and if prolonged sufficiently, will die. In 28 rats brought to this stage of recovery, all were weaned to sweet fluid and all ate dry food to maintain their weight in the process. For every animal, removal of the sweet fluid led to dehydration because it did not accept water and, consequently, stopped eating. Restore the sweet fluid, however, and it is accepted immediately (as seen in Figure 1). They eat dry food shortly thereafter and then regain their weight. In this stage, therefore, the animals do not eat because they do not drink.

The counter-experiment was performed on four rats. The saccharin solution was again removed, but water, its essential component, was supplied by hydrating the animals intragastrically with 2 to 3 milliliters every hour automatically through a chronically implanted stomach tube. As seen in Figure 1 in the hatched area labeled 2 starting on the 64th postoperative day, the animals continue to eat dry food normally when hydrated, even though they no longer have saccharin to drink. Hydration either by drinking sweet solutions or by intragastric injections allows normal feeding. It is clear, then, that in this stage of the lateral hypothalamic syndrome, the refusal to drink water produces dehydration which in turn prevents feeding.

Does dehydration also account for the failure to eat in Stage I? Will these animals eat immediately after operation if they are hydrated? Chronic gastric tubes were inserted in five animals before they were subjected to bilateral lateral hypothalamic lesions. From the time they recovered from the anesthetic, usually the day after operation, they were hydrated

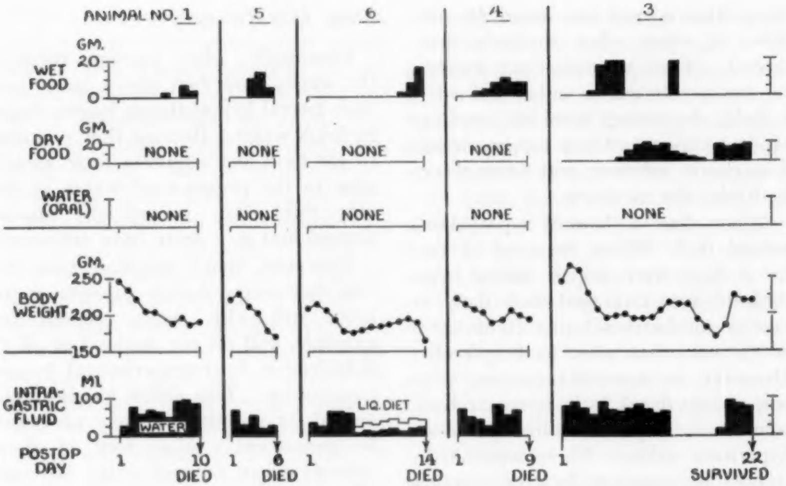


FIG. 2. Food and water intake and body weight of five animals with lateral hypothalamic lesions hydrated automatically from the time of operation. (Four of the animals died after a period of complete aphagia and anorexia despite hydration.)

with 2 to 3 milliliters of water automatically every hour. Figure 2 shows the results. All animals showed complete, immediate postoperative aphagia, varying in duration from 2 to 10 days. Then they all displayed anorexia rather than aphagia, eating some wet food but not enough to maintain their weight. Four of the five animals died. The fifth began to eat dry food on the 7th day and was able to maintain its weight. When hydration was stopped, it stopped eating and lost weight; it ate again immediately after hydration was resumed. It had, therefore, recovered from its initial aphagia and anorexia and passed into Stage III, where it ate if, and only if, it was hydrated. Because all of the animals were initially aphagic and because four of the five animals died despite adequate hydration, we conclude that intragastric hydration initiated in the immediate postoperative period does not assure normal feeding (Morrison & Mayer, 1957). The aphagia and anorexia of Stages I and II are not due just to dehydration.

Why do the rats in Stage III after lateral hypothalamic lesions refuse tap or distilled water and yet readily accept saccharin solutions? Do they drink the sweet saccharin solution because they are thirsty or do they take it in the same way that they take a liquid diet—as a food? A normal animal increases its intake of saccharin solution both when deprived of food (Carper & Polliard, 1953) and when dehydrated. Four normal rats, deprived of food for 3 days, took little or no water but drank increasing amounts of saccharin (from two to four times as much as they did when feeding normally). Three normal rats, dehydrated by an intraperitoneal injection of 2 milliliters of 1 M NaCl per 100 grams body weight (Adolph, Barker, & Hoy, 1954), showed an average intake of 14 milliliters of 0.2% saccharin solution in 2 hours, whereas

three other normal rats drank 16 milliliters of water when similarly dehydrated. Thus a normal rat treats a saccharin solution as either a food or a fluid, depending upon its condition of deprivation. When hungry it eats a saccharin solution, and when thirsty it drinks the saccharin.

What does a lateral hypothalamic animal do? When deprived of food for 3 days, four adipsic lateral hypothalamic rats increased their daily intake of saccharin solution (from 20 to 300% more than when feeding freely). However, on repeated occasions, seven adipsic rats drank little or no saccharin solution in response to intraperitoneal hypertonic saline. There were 20 instances of complete lack of response for as long as 5 hours and five instances when they drank only 2 or 3 milliliters. Even a large dose of salt that proved lethal failed to elicit increased fluid intake. Because lateral hypothalamic animals do not increase their saccharin intake when dehydrated but do when deprived of food, they are, in effect, eating the saccharin, not drinking it. They are hungry, not thirsty. Because they do not respond in this stage to levels of dehydration that produce vigorous drinking in normal animals, their refusal to drink water in all three initial stages of recovery may be due to a severe impairment of the central mechanism for water intake regulation.

The refusal to drink water may be permanent, as illustrated in Figure 1. The animal died 200 days after operation, never again having drunk water. It sustained itself for most of that time by "eating" sweet fluid and eating dry food. Three other animals showed permanent adipsia until they died after 267, 230, and 107 days, respectively. One animal, still living, has been adipsic for 485 days.

Stage IV: Recovery

Eventually, after passing through the stages described above, most rats with lateral hypothalamic lesions begin to drink water. Because they continue to eat food and maintain their weight now in the presence of water as the only fluid, they superficially appear normal and in a sense have recovered.

However, more sensitive tests reveal that severe deficits in water regulation still exist. Some animals, for example, still do not respond at all to dehydration by intraperitoneal hypertonic saline. This deficit has been observed for months and may very well be permanent. Also, four of these animals when offered water for only 2 hours each morning did not compensate by taking their daily requirement (approximately 25 milliliters) in that time as normal animals do. They took little or nothing, became progressively dehydrated, stopped eating, and lost weight until they were nearly dead. If they drank at all, they did so shortly after the water was presented. They typically drank 1 or 2 milliliters (never more than 5 milliliters), moved to the back of the cage and slept. When returned to 24-hour access to water, they again managed in an apparently normal manner.

There is a curious paradox here. These animals *do* drink when the water is available overnight but do *not* drink in response to rapid dehydration with hypertonic salt nor do they drink enough when water is available for 2 hours in the morning. The reason for this apparent contradiction was discovered by observing what the animals were doing when they drank. By attaching a drinkometer (Stellar & Hill, 1952) to the water tube and recording the pattern of water ingestion, we found that these animals drank nothing during the day, but drank in

short bursts of rapid licking almost continuously throughout the entire night. This is when they normally eat. In addition, if offered both food and water simultaneously at night, they began to drink only after they had begun to eat. Shortly after eating a little dry food, they drank a little water, ate more dry food, drank more water, and so on throughout the night. One of the animals showed this behavior in its most extreme form. It would stand near the drinking tube holding a pellet in its front paws, munch on the pellet for a few minutes, lick actively at the water, and then go back to chewing on the pellet. Its eating was extremely inefficient. It would bite large chunks off each pellet, chew them a little and then drop them while it drank a small quantity of water. In the course of the night it would waste as much if not more food than it actually ate by dropping small chunks of it through the mesh floor of its cage. Its drinking was also abnormal. Only small amounts of water were drunk during bursts of active and precise licking. Although 25 to 30 milliliters of water were consumed, the cumulative record showed continuous licking throughout the night, interrupted by pauses for short bouts of eating. The animal licked all night to consume an amount of water that a normal rat could drink in several sustained draughts. It was as if the animal were drinking not to quench its thirst but simply to wet its mouth.

These observations suggest that such an animal drinks only while it eats. To demonstrate this, it was deprived of water for 24 hours and offered access to it for 2 hours during the night (8 to 10 P.M.) when it normally ate. It drank 22 milliliters of water while eating 12 grams of food. The amount of water deprivation made no difference. It drank only in proportion to the amount it ate. After 48 hours of water deprivation it drank 22 milliliters of water while eating 11 grams of food. When the experiment was repeated in the absence of food so the animal could not eat, it again did not drink any water in 2 hours despite as much as 48 hours of water deprivation. As soon as the food was replaced, the animal started eating and only after eating for a while did it start to drink. The same result was obtained in all four animals that had previously failed to drink adequately in the morning. The animals were then trained to shift their eating into the morning hours (10 A.M. to 12 noon) by offering them food only at that time. Now that they were eating in the morning, they drank enough water in 2 hours to maintain their weight on this regime for several days. Because the animals drink only while they eat, we call this phenomenon "prandial drinking."

In the same way that the anorexic animals seen in Stage II are eating without caloric regulation, so in Stage IV, when the animals first begin to drink water, they are drinking without response to bodily dehydration. Just as palatability was sufficient to arouse appetite for food, so the mouth factors involved in eating dry food may be sufficient to arouse prandial drinking.

These animals are greatly affected by slight changes in the taste of their water. Normal rats will drink bitter quinine hydrochloride solutions, up to a concentration of 1% w/v, if they are the only fluids available. When ingested at this concentration, 15 to 20 milliliters of quinine hydrochloride is lethal to a rat. A normal rat, therefore, will drink quinine-treated water until it becomes a poison. As this point is approached, they drink less

and less, gradually stop eating, and die. Lateral hypothalamic rats that have been drinking water reject quinine hydrochloride solutions at a concentration as low as 0.005%. They refuse it day after day, stop eating, lose weight, and die. But when they are hydrated intragastrically with it, they thrive. Thus, the refusal to eat and drink until death seen in the earlier stages of the lateral hypothalamic syndrome can be reinstated in apparently recovered animals merely by making their water slightly bitter.

A normal rat, as it becomes thirstier, overcomes the barrier to ingestion imposed by the bitter taste of the quinine up to the lethal concentration and drinks enough to stay alive. But a rat with lateral hypothalamic damage does not. Even though it drinks normal amounts and maintains itself when water is constantly available, the lateral hypothalamic animal fails to overcome the quinine taste barrier. In addition, it fails to drink adequate quantities when access to water is limited despite prolonged dehydration, and fails to respond to intraperitoneal saline challenge. These deficits in motivation and in water regulation imply that the central neural thirst mechanism is still severely impaired. The powerful urge to drink is still absent. These animals drink only while they eat, *not* in response to dehydration. There is a striking resemblance here to Adolph's (1957) description of the ontogeny of food and water regulation. The infant rat develops hunger before thirst and, in the neonatal period, does not respond to dehydration. It is adipsic but not aphagic.

Although, according to many measures, food intake has been normal since early in Stage III, more sensitive tests again reveal continuing deficits. Here, too, taste seems to be an important factor. Even though they are apparently eating normally, as little as 0.05% quinine hydrochloride w/v, mixed thoroughly in dry Purina powder, is often sufficient to cause rats with lateral hypothalamic lesions to refuse food completely to the point of extreme starvation. Normal rats readily accept food with up to 1.25% quinine, at which point they begin to decrease their intake and lose weight. When faced with the bitter taste of quinine, rats with lateral hypothalamic lesions, therefore, display a greater aversion to a negative taste in food. Palatability is again a powerful determinant of food intake in these animals. In the anorexic stage (Stage II), only highly palatable foods induced them to eat. And in the recovery stage, a slight negative taste prevents them from eating. This resembles the finicky eating observed in obese hyperphagic rats after more medial hypothalamic lesions (Teitelbaum, 1955). This finickiness implies either an impaired motivation for food (Miller, Bailey, & Stevenson, 1950; Teitelbaum, 1957) or an increased reactivity to its taste; presumably it is present in its most extreme form in the early stages of the lateral hypothalamic syndrome.

The recovery process continues even after the animal has started drinking water again. Some rats become more responsive to intraperitoneal injection of hypertonic saline and indeed three of seven rats tested eventually responded as well as a normal animal. Similarly in these same three animals it required as much as 0.1% quinine to prevent them from drinking whereas earlier they had been stopped by a 0.005 or 0.01% quinine solution. A gradual recovery can also be seen in the gross regulation of water intake, food intake, and weight. Immediately after entering the recovery stage, and for some time thereafter, some rats

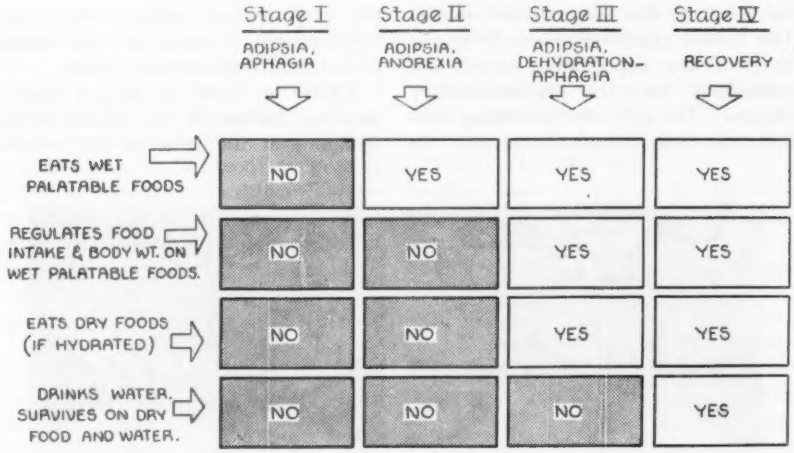


FIG. 3. Stages of recovery seen in the lateral hypothalamic syndrome. (The critical behavioral events which define the stages are listed on the left.)

with lateral hypothalamic lesions drink water in less than normal amounts, eat correspondingly less, and so maintain their weights at a subnormal level (210 to 220 grams). This hypodipsia may last for months, but usually it disappears. Such a rat begins to drink and eat more, and it regulates its weight at 270 to 280 grams, a more normal weight for its age.

The four stages of recovery from lateral hypothalamic lesions are summarized in Figure 3, which lists the critical behavioral events whose presence or absence defines those stages. First, there is a stage of "aphagia and adipsia" in which the animal completely refuses to eat or drink; second, "anorexia and adipsia" in which it eats, but not enough to maintain its weight, and does not drink; third, "adipsia and dehydration-aphagia," in in which the animal fails to eat only because it does not drink; and finally the "recovery" stage in which the animal drinks water, eats dry food, and will survive, but still shows certain deficits in eating and drinking behavior. It must be emphasized that these stages in recovery will not be observed unless the animal is offered wet and palatable foods from the time of the operation. If only water and dry food are available, the rat will die unless it rapidly reaches Stage IV, at which point it accepts water and can, therefore, eat dry food.

ANATOMICAL ANALYSIS

The duration of these stages is variable and seems to depend on both the locus of the lesions and the amount of tissue damaged. The lesions in 13 animals were studied microscopically using thionin-stained celloidin sections through the diencephalon. Tissue in the lateral hypothalamic areas was destroyed bilaterally in all animals. Typically, the lesions were large, extending medially to the mammillothalamic tracts and the fornices (the fornix itself was only rarely destroyed) and laterally into the medial portions of the internal capsules and optic tracts. Damage was greatest at

the level of the ventromedial nuclei. The lesions extended anteriorly to the level of the supra-optic nuclei and posteriorly into the pre-mammillary regions. Dorsally, the overlying subthalamus was always involved. In the majority of animals, the most ventromedial portions of the lateral hypothalamic areas were spared.

Figure 4 shows a frontal section showing lesions in the lateral hypothalamus at the level of the ventro-

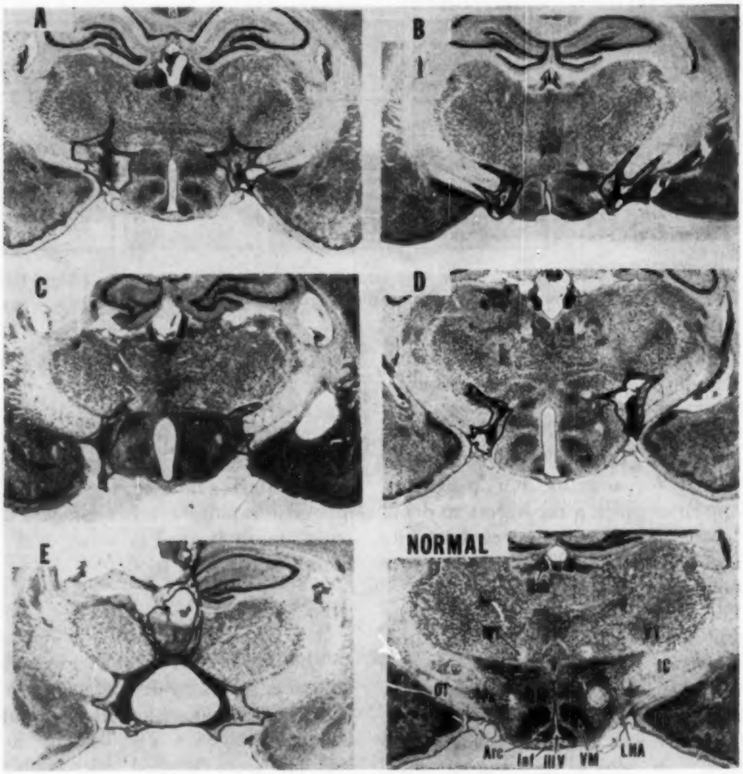


FIG. 4. Photomicrographs of thionine-stained sections through the tuberal region of the hypothalamus showing lesions in five animals with interesting variations of the lateral hypothalamic syndrome. (Damaged tissue is outlined in black. A: Large, symmetrical lesions producing permanent adipsia. Recovery graphed in Figure 1. B: Transient aphagia and anorexia, despite immediate postoperative hydration, followed by recovery of food and water intake. Animal 3 in Figure 2. C: Died on sixth postoperative day despite adequate hydration. Showed aphagia and anorexia. Animal 5 in Figure 2. D: Asymmetrical lesions producing transient adipsia without aphagia. E: Massive lesions involving both medial and lateral tissue producing hypothalamic hyperphagia and permanent adipsia.

Normal animal: Arc, arcuate nucleus; DM, dorsomedial nucleus; Fx, fornix; Hab, habenula; IC, internal capsule; Inf. infundibulum; LHA, lateral hypothalamic area; MT, mammillothalamic tract; OT, optic tract; Rh, rhinencephalon; VM, ventromedial nucleus; VT, ventral thalamus; ZI, zona incerta; III V, third ventricle.)

medial nuclei in five animals who exhibited interesting variations of the lateral hypothalamic syndrome. The extent of damage has been outlined heavily in black. With the very large lateral hypothalamic lesions of Section A, each stage in the recovery process lasted for weeks or months. This is the animal whose recovery is shown graphically in Figure 1. This rat showed prolonged aphagia and anorexia (38 days), and it never drank water again. Such an animal, because of the dehydration-aphagia that is present when it refuses to drink water, would seem to be permanently aphagic on a dry diet. It was only by inducing it to accept a palatable, wet diet that the stage of pure adipsia could be demonstrated. With smaller or asymmetrical lesions, as in Sections B, C, and D, the entire sequence passes very quickly, and the animal may begin the recovery not at Stage I, but at any of the later stages. These animals show aphagia and anorexia for only a few days or a week if offered a palatable liquid diet, and recover water intake within 1 or 2 months. Section B is from the brain of animal 3 in Figure 2. It refused to eat for 3 days despite automatic hydration, then nibbled wet food for 4 days, but remained adipsic for 1 month after operation. Section C shows the lesions of Animal 5 in Figure 2. It was aphagic and anorexic until it died even though hydrated. No attempt was made to keep this animal alive by tube-feeding, but since it had already begun to eat appreciable amounts of wet food, it probably would have regulated its caloric intake within a few more days. Section D shows lateral hypothalamic lesions which produced only adipsia, without any aphagia or anorexia. The animal accepted wet mash or liquid diet immediately and regulated its weight, but it refused to drink water and showed dehydration-aphagia on a dry diet. This effect closely resembles the earlier findings of adipsia without aphagia in the rat by Montemurro and Stevenson (1957) and in the dog by Andersson and McCann (1956). Finally, Section E, at a slightly more rostral level, shows a huge lesion that involves the entire hypothalamus except for the infundibulum, destroying both the ventromedial nuclei and the lateral hypothalamic areas, and extending laterally well into the internal capsules. This animal was aphagic and anorexic for only 4 days, then regulated its intake of wet food. It refused to drink water until it died 107 days later. When offered a wet diet it overate and became obese; but every time it was offered a dry diet and water, it refused to drink, stopped eating as a result of dehydration, and lost weight precipitously. This verifies the earlier findings of Williams and Teitelbaum (1959), who showed that hypothalamic hyperphagia can coexist with the lateral hypothalamic syndrome. It cannot be manifested in Stages I or II during which caloric regulation is not seen, but in Stage III, when caloric regulation appears, so does hypothalamic hyperphagia. But hydration is necessary. With a dry diet and water the only fluid available, the animal does not drink, and hypothalamic hyperphagia is masked by adipsia. In addition, such animals make it clear that damage to the ventromedial regions of the hypothalamus can produce hyperphagia even when the lateral hypothalamic areas are largely destroyed.

In general, the severity of the deficits was correlated with the size of the lesions. Transient deficits and rapid recovery followed small and asymmetrical lesions, whereas large lesions permitted only slow and incomplete recovery through stages of

long duration. This suggested that the recovery of function was mediated by the undamaged tissue around the lesions. Additional lesions in the tissue immediately adjacent to the original damage were made in six recovered animals. In all of them, additional lesions were found to reinstate the syndrome. After appropriate lesions the reoperated animals became aphagic, anorexic, and adipsic all over again, and the recovery followed the same general pattern. The phenomenon was particularly clear in one animal. The recovery process was repeated four times. Each successive recovery, in this animal, was slower. The fourth recovery proceeded only to anorexia and was terminated by a fatal gastric hemorrhage. Other animals have been rendered permanently anorexic by additional lesions, in one for as long as 270 days. Therefore, the undamaged tissue surrounding the original lesion participates in the recovery. In addition, this result suggests that the experience gained in the initial recovery from lateral hypothalamic lesions did not protect the animal from the effects of additional damage.

Since the lateral hypothalamic syndrome in rats is the result of a complex interaction of deficits in both feeding and drinking, the lateral hypothalamus must contain tissues concerned with both functions. The structures involved could be cells within the lateral hypothalamus or fibers of passage from other areas. Morgane (1961a) reports a syndrome of aphagia and adipsia in the rat with lesions in the apex of the *globus pallidus*, completely outside the lateral hypothalamus. He points out that the major efferent pathways from the *globus pallidus* (*ansa* and *fasciculus lenticularis*) are usually interrupted by lateral hypothalamic lesions. Thus, extrahypothalamic structures participate in the regulation of food and water intake.

Whatever the nature of the neuroanatomical systems involved in feeding and drinking, some parts of them must overlap in the lateral hypothalamus. Is it possible to separate them? A chemical separation of the systems has been demonstrated in the rat by Grossman (1960). He showed that from the same site within the perifornical region of the hypothalamus, cholinergic drugs elicit drinking and adrenergic drugs produce feeding. This implies that the functions of each of the systems may be dependent upon different neurohumoral agents.

An anatomical separation may be possible because small, asymmetrical lesions produce adipsia without aphagia and pure effects have been produced by electrical stimulation. Drinking without feeding has been observed in the goat and rat (Andersson & McCann, 1955–56; Greer, 1955), as well as feeding without drinking (Larsson, 1954; Miller, 1957; Smith, 1956). Thus, it is quite possible that two systems which do not entirely overlap, one concerned with drinking and the other with feeding, merge in the lateral hypothalamus, and so destruction there yields the combination of deficits in feeding and drinking behavior that makes up the lateral hypothalamic syndrome.

Morgane (1961b) has suggested a further subdivision of the lateral hypothalamic feeding system. He finds that rats with lesions in what he has called the "far-lateral" region of the lateral hypothalamus (centered just medial to the internal capsule) lose weight more rapidly while suffering aphagia and adipsia and are less likely to recover than animals with lesions in the "mid-lateral" hypothalamus (centered just lateral to the fornix). According to Morgane, recovery is

the rule following mid-lateral lesions. But recovery is the exception, if not an impossibility, after far-lateral lesions. He finds it difficult to maintain the far-lateral animal alive on a tube-feeding regime that is adequate for the mid-lateral animal. He suggests that the far-lateral region of the hypothalamus contains pallidofugal structures (*ansa* and *fasciculus lenticularis*) that are essential for the basic feeding reflexes. Lesions here should therefore produce permanent aphagia. In addition, the failure of the far-lateral animal to survive on tube-feedings leads Morgane to attibute its rapid decline to a form of "metabolic failure."

In our own experience, as described in this paper, many of the animals that have shown the typical stages of recovery were found to have extensive damage in the far-lateral hypothalamus (see Figure 4, especially Section A). Clearly these animals have very severe and prolonged deficits. In this we agree with Morgane. However, they do feed again, although they may remain adipsic. In the early stages of their recovery, they can be maintained on tube feeding if cared for assiduously. The more rapid decline in weight seen in similar animals and attributed to metabolic failure by Morgane can not, therefore, be due to an inability to utilize food but may be related to changes in activity, temperature regulation, or metabolic rate. Stevenson[6] has shown that lesions within the lateral hypothalamus may produce an increase in metabolic rate and has suggested that this increased energy expenditure may account for the rapid postoperative decline of the far-lateral animal. In addition, we must re-emphasize that lateral hypothalamic lesions produce deficits in *both* feeding and drinking. Prolonged deficits then may simply represent the summated effects of extensive destruction of both systems rather than selective involvement of "more basic" and functionally distinct portions of the feeding system alone.

[6] Stevenson, J. A. F. Personal communication, 1961.

SUMMARY AND CONCLUSIONS

Although we have made interpretations throughout this paper as new phenomena were described, there are several generalizations from this work that deserve emphasis. Each of them has changed our thinking about the lateral hypothalamic control of feeding and drinking.

We have confirmed the idea that there are tissues in the lateral hypothalamus that are concerned with both feeding and drinking. Aphagia is always accompanied by adipsia. We have never seen pure aphagia—i.e., absence of feeding where drinking is unimpaired. Any impairment in feeding then that follows these lesions can only be fully understood as the combined result of deficits in both regulations.

Recovery is the rule following these lesions, if the animals are maintained long enough. Barring accidents, or gastrointestinal pathology, it is possible to keep these animals alive indefinitely. Although we have seen prolonged deficits, we have never seen an animal with lateral hypothalamic lesions that did not show some recovery.

Feeding recovers before drinking. The quality of the diet is crucial for this observation. At first the animal will eat only wet and palatable foods. But if it is offered nothing except dry pellets and water, it will not eat until after it has begun to drink, and this relationship will be misunderstood. It appears to us, therefore, that lateral

hypothalamic lesions impair feeding less severely than drinking.

Eating does not necessarily imply regulation. When the animals first begin to eat they are anorexic. They eat only highly palatable foods, and do not eat enough to regulate caloric intake or body weight. Their feeding seems to be evoked by the taste and smell of the food. They are drawn to it by appetite, not driven to it by hunger.

Hunger returns, often abruptly. A new control reappears, enabling the animal to regulate its caloric intake. The central feeding mechanism is now responsive to some variable associated with caloric need. If this variable were identified, we would know one control system that is sufficient for caloric regulation in the normal rat.

Even after the return of regulatory feeding, adipsia persists. It can mask caloric regulation if only dry food and water are available. The animal will not drink water and aphagia will be the result of dehydration. If the animal is induced to eat saccharin solutions, it hydrates itself involuntarily and then eats dry pellets. Thus, adipsia can occur without aphagia.

Adipsia can be permanent. Some animals never drink water again, though they live out their natural lives. In our experience, aphagia is always transient. We have never seen permanent aphagia after lateral hypothalamic lesions. This confirms our belief that the central neural tissue regulating thirst is more completely impaired by these lesions than is the tissue controlling feeding.

Drinking frequently reappears. Initially the animals drink only while they eat. Just as anorexia is feeding without response to starvation so prandial drinking is drinking without response to dehydration. Here again, ingestion does not necessarily imply regulation. Like hunger, thirst may eventually reappear. The central thirst mechanism is now responsive to dehydration.

Although regulation of food and water intake has recovered, motivation is still impaired. The recovered animal is greatly affected by the taste of both its food and water. Addition of small amounts of quinine deters it from eating and drinking. Wherever taste is excessively important, motivation is impaired.

Finally, in the lateral hypothalamic syndrome, the initial transition from aphagia to anorexia and from adipsia to prandial drinking seems to depend upon recovery of the influence of peripheral sensations (taste and smell of food, or dry mouth). The next transition involves the recovery of central neural controls permitting not only ingestion but ingestion with regulation. Anorexia gives way to hunger and prandial drinking to true thirst. This supports the commonly accepted view that feeding and drinking are both under multiple control (Brobeck, 1960; Stellar, 1954). Because all controls may operate simultaneously in the normal animal, it has been difficult to identify those essential for regulation. In the lateral hypothalamic syndrome, the reappearance of these controls is separated in time during recovery. They are not all operating simultaneously. Controls that operate in the later stages of recovery are absent in earlier stages. It should be possible to identify both the controls that are present and those that are absent during any particular stage by subjecting the animal in that stage to the variables known to influence feeding and drinking in the normal animal. If we find controls that are absent before regulation, and which

reappear with it, we may then be able to specify the controls necessary for regulation.

REFERENCES

ADOLPH, E. F. Ontogeny of physiological regulations in the rat. *Quart. Rev. Biol.*, 1957, **32**, 89–137.

ADOLPH, E. F., BARKER, J. P., & HOY, P. A. Multiple factors in thirst. *Amer. J. Physiol.*, 1954, **178**, 538–562.

ANAND, B. K., & BROBECK, J. R. Hypothalamic control of food intake. *Yale J. Biol. Med.*, 1951, **24**, 123–140.

ANAND, B. K., DUA, S., & SCHOENBERG, K. Hypothalamic control of food intake in cats and monkeys. *J. Physiol.*, 1955, **127**, 143–152.

ANDERSSON, B., & MCCANN, S. M. Drinking, antidiuresis and milk ejection from electrical stimulation within the hypothalamus of the goat. *Acta physiol. Scand.*, 1955–56, **35**, 191–201.

ANDERSSON, B., & MCCANN, S. M. The effect of hypothalamic lesions on the water intake of the dog. *Acta physiol. Scand.*, 1956, **35**, 312–320.

BROBECK, J. R. Regulation of feeding and drinking. In J. Field (Ed.), *Handbook of physiology. Sect. 2, Vol. 2.* Baltimore: Williams & Wilkins, 1960.

BROBECK, J. R., TEPPERMAN, J., & LONG, C. N. H. Experimental hypothalamic hyperphagia in the albino rat. *Yale J. Biol. Med.*, 1943, **15**, 831–853.

BRÜGGER, M. Fresstrieb als hypothalamisches Symptom. *Helv. physiol. pharmacol. Acta*, 1943, **1**, 183–198.

CARPER, J. W., & POLLIARD, F. A comparison of the intake of glucose and saccharin solutions under conditions of caloric need. *Amer. J. Psychol.*, 1953, **66**, 479–482.

EPSTEIN, A. N. Reciprocal changes in feeding behavior produced by intrahypothalamic chemical injections. *Amer. J. Physiol.*, 1960, **199**, 969–974. (a)

EPSTEIN, A. N. Water intake without the act of drinking. *Science*, 1960, **131**, 497–498. (b)

EPSTEIN, A. N., & TEITELBAUM, P. Regulation of food intake in the absence of taste, smell, and other oro-pharyngeal sensations. *J. comp. physiol. Psychol.*, in press.

FRENCH, J. D., PORTER, R. W., VON AMERONGEN, F. K., & RANEY, R. B. Gastrointestinal hemorrhage and ulceration associated with intracranial lesions. *Surgery*, 1952, **32**, 395–407.

GREER, M. A. Suggestive evidence of a primary "drinking center" in hypothalamus of the rat. *Proc. Soc. Exp. Biol. Med.*, 1955, **89**, 59–62.

GROSSMAN, S. P. Eating or drinking elicited by direct adrenergic or cholinergic stimulation of hypothalamus. *Science*, 1960, **132**, 301–302.

LARSSON, S. On the hypothalamic organization of the nervous mechanism regulating food intake. *Acta physiol. Scand.*, 1954, **32**, Suppl. No. 115, 7–40.

MILLER, N. E. Experiments on motivation. *Science*, 1957, **126**, 1271–1278.

MILLER, N. E., BAILEY, C. J., & STEVENSON, J. A. F. Decreased "hunger" but increased food intake resulting from hypothalamic lesions. *Science*, 1950, **112**, 256–259.

MONTEMURRO, D. G., & STEVENSON, J. A. F. Adipsia produced by hypothalamic lesions in the rat. *Canad. J. Biochem. Physiol.*, 1957, **35**, 31–37.

MORGANE, P. J. Alterations in feeding and drinking behavior of rats with lesions in globi pallidi. *Amer. J. Physiol.*, 1961, **201**, 420–428. (a)

MORGANE, P. J. Medial forebrain bundle and "feeding centers" of the hypothalamus. *J. comp. Neurol.*, 1961, **117**, 1–26. (b)

MORRISON, S. D., & MAYER, J. Adipsia and aphagia in rats after lateral subthalamic lesions. *Amer. J. Physiol.*, 1957, **191**, 248–254.

SMITH, O. A. Stimulation of lateral and medial hypothalamus and food intake in the rat. *Anat. Rec.*, 1956, **124**, 363–364.

STELLAR, E. The physiology of motivation. *Psychol. Rev.*, 1954, **61**, 5–22.

STELLAR, E., & HILL, J. H. The rat's rate of drinking as a function of water deprivation. *J. comp. physiol. Psychol.*, 1952, **45**, 96–102.

STELLAR, E., & KRAUSE, N. P. New stereotaxic instrument for use with the rat. *Science*, 1954, **120**, 664–666.

STEVENSON, J. A. F., & RIXON, R. H. Environmental temperature and deprivation of food and water on the spontaneous activity of rats. *Yale J. Biol. Med.*, 1957, **29**, 575–584.

TEITELBAUM, P. Sensory control of hypothalamic hyperphagia. *J. comp. physiol. Psychol.*, 1955, **48**, 156–163.

TEITELBAUM, P. Random and food-directed activity in hyperphagic and normal rats.

J. comp. physiol. Psychol., 1957, **50**, 486–490.

TEITELBAUM, P., & STELLAR, E. Recovery from the failure to eat produced by hypothalamic lesions. *Science*, 1954, **120**, 894–895.

WHEATLEY, M. D. The hypothalamus and affective behavior in cats. *Arch. Neurol. Psychiat., Chicago*, 1944, **52**, 296–316.

WILLIAMS, D. R., & TEITELBAUM, P. Some observations on the starvation resulting from lateral hypothalamic lesions. *J. comp. physiol. Psychol.*, 1959, **52**, 458–465.

WYRWICKA, W., & DOBRZECKA, C. Relationship between feeding and satiation centers of the hypothalamus. *Science*, 1960, **132**, 805–806.

(Early publication received August 14, 1961)

Psychological Review
1962, Vol. 69, No. 2, 91–108

A STATISTICAL THEORY OF DYNAMIC CONTOUR PERCEPTION[1]

W. M. SMITH
Dartmouth College

AND

W. L. GULICK
University of Delaware

Bartley's (1941) statement made almost twenty years ago, ". . . that the matter of contour formation is one of the central phenomena of sensory psychology and nerve physiology" (p. 334), certainly is as true today as it was when he wrote it. Indeed, the way in which borders are perceived is a fundamental problem in visual science.

It is very clear that the classical idea of the retinal mosaic as a system of insulated receptors cannot meet adequately various facts of visual acuity, brightness discrimination, and contour definition. Visual acuity is greater than the classical notion implies, and contour perception is not dependent upon two parallel arrays of receptors, one array stimulated, the other not. Inadequate consideration of postretinal structures has been the major failing of what is now considered to be the classical view of visual acuity.

The work of Marshall and Talbot (1942) is a formulation of a theory of visual acuity and contour resolution in terms of the structure and the function of retinal *and postretinal* neural elements. Drawing upon recent findings and concepts of neurophysiology, they view such basic processes of acuity and contour perception as dependent upon the interrelated activity of *populations* of receptors and neural elements. Appropriately, their formulation now is referred to as a statistical theory of visual function.

Although we shall not present a review of the Marshall and Talbot work, for expository purposes later it will be helpful to mention some general features of their theory here. Their analysis rests heavily upon the neurological concepts of reciprocal overlap, lateral and vertical summation in the visual system, physiological nystagmus, and the nature of the neural recovery cycle. Especially important is the fact that their theory places the locus of the critical aspects of contour perception in the visual cortex rather than at the retinal level.

In their original work Marshall and Talbot were concerned primarily with the general question of sensory acuity, but in more recent years other investigations have extended the statistical theory to other aspects of vision. Osgood and Heyer (1952), for example, have applied it to the analysis and interpretation of figural aftereffects and have claimed the statistical theory to be superior to the field conception offered by Köhler and Wallach (1944). Regarding figural aftereffects, however, both the field and the statistical interpretations have had their critics (Deutch, 1956; Smith, 1952). More recently, Day (1956, 1957) has used statistical theory in the interpretation of phenomenological reports and experimental data concerning form perception. In particular, he has employed statistical theory in considering the question of the emergence of form in the peripheral retina and in interpreting data from studies of area and area-perimeter ratios in form perception. Finally, Osgood (1953, pp. 245–47) has pre-

[1] This work has been supported by the National Science Foundation.

sented a brief account of how statistical theory can be applied to the problem of apparent movement.

It is with the perception of movement that the present paper is concerned. We shall attempt to extend statistical theory to a treatment of perceived movement. First, we shall describe briefly the basic experimental situation. Then we shall summarize the results from which the theoretical treatment was developed. Finally, we shall present an account of our modification and extension of statistical theory with supporting evidence.

Dynamic Contour Perception

For present purposes we define "dynamic contour perception" as the perception of the borders of moving stimuli. The definition could be modified to include instances wherein the observer moves rather than the stimulus. We shall not consider this modification here.

In two previous papers (Smith & Gulick, 1956, 1957) data have shown that the perception of the borders of moving stimuli is influenced not only by the velocity of the stimulus, but also by the length of time the stimulus is observed in a fixed position prior to movement. The results of these experiments indicated that the border of a small square could not be seen clearly at velocities exceeding about 13 degrees of visual angle per second. However, when the stimulus was presented in a fixed position prior to movement, the maximum velocity at which the edges could be perceived was increased by a factor or two or more in most instances. A systematic study of the combinations of velocity and initial exposures needed to maintain contour revealed a linear relationship up to exposures of about 350 milliseconds and velocities of about 30 degrees of visual angle per second. For every increase in velocity that produced loss of contour, an increment in the duration of the initial exposure could be utilized to regain it. In more recent unpublished studies in which the duration of exposure of the initial stationary phase of the stimulus has been extended to 1 second, we find that the ability to see an edge with increasing velocities of movement is related curvilinearly to the duration of the initial stationary phase. At velocities exceeding about 40 degrees per second it appears that the fixed phase of presentation, regardless of its duration, will not enable one to see a moving edge clearly. In an earlier paper (Smith & Gulick, 1957) we offered a preliminary statement of statistical theory as applied to the combined effects of the initial exposure of the stimulus in the fixed position and the movement phase. In the present paper we wish to expand this preliminary theoretical treatment.

Historical Background of the Model

When an observer's total visual field consists of a light-dark pattern divided vertically, there is no apparent gradient of gray where light meets dark. The discontinuity is clearly defined. How the visual system is able to "construct" an edge from such a distribution of light energy striking visual receptors is not well understood. However, it is known that the abrupt gradient of reflected light which exists at the border of a stimulus object is not present once the pattern of light reaches the retina (Fry, 1955). The abrupt gradient becomes less abrupt so that the intensity of light comprising the geometric image of a border at the retina approximates the integral of a normal distribution. It is this pattern of light energy which courses back and forth

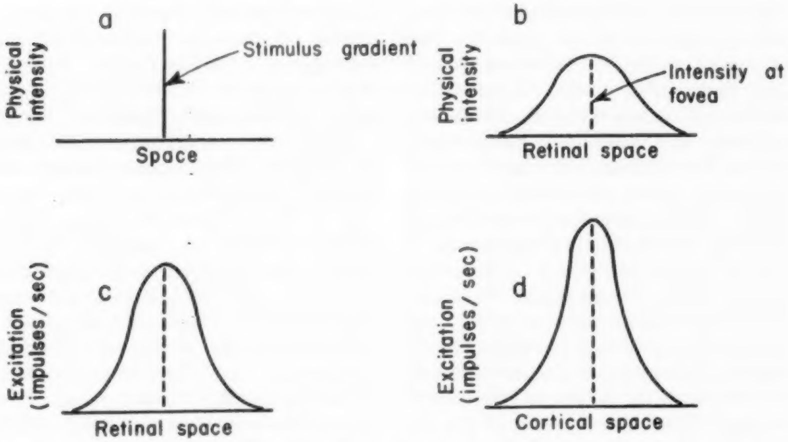


FIG. 1. Hypothetical distributions illustrating transformation of stimulus from environment to cortical space. (The platykurtic distribution of light at the fovea, b, is transformed to a more leptokurtic distribution of excitation within the afferent system, c. At the cortical level, d, the distribution of excitation is peaked further and is the direct correlate of perceived contour.)

across a population of receptors as the eye undergoes nystagmic movements.

If, instead of a single border, an observer is shown a pattern of two parallel borders, such as a bright, thin vertical line (near threshold value), the distribution of light of this pattern on the retina approximates a normal curve. As the angular width of such a pattern increased, the symmetrical distribution at the retina would be changed progressively, becoming flattened on top. Each border, however, would be represented on the retina by a distribution of light similar to that for the discontinuity of the vertically divided field mentioned previously.

It is the central proposition of statistical theory that the *cortical* representation of a perceived edge is a symmetrical distribution of excitation in Area 17. The leptokurtic distribution of excitation is correlated with the abrupt change in intensity at the stimulus boundary, but only by virtue of a very complicated series of events and mechanisms intervening between cornea and cortex. The abrupt change of energy level in the stimulus is not reproduced at the retina because of distortion introduced by passage through ocular media. Slight irregularities of the surface of the eye, errors of refraction and of aberration, and entopic stray light all tend to reduce the abruptness of the stimulus gradient as it is projected on the retina.

In Figure 1 we have attempted to show schematically the transformation from stimulus gradient to cortical distribution. A line in the visual field presents an abrupt gradient of energy. At the retina the distribution of energy assumes the platykurtic form as shown in Figure 1b. The relative flatness of this distribution is due to the scattering of light by the distorting factors noted previously, plus the effects of physiological nystagmus. This distribution of light represents the prox-

imal stimulus which triggers the primary receptors. If one plots the rate of neural firing for different retinal loci, then a distribution of excitation results. This hypothetical distribution is shown in Figure 1c. The peaking of this distribution is produced by the combined effects of numerous mechanisms. First, neural elements located near the center of the distribution in Figure 1b are stimulated by the most intense light. Accordingly, the number of fibers stimulated as well as the average rate of firing of these neurons is greater than in those neural elements located near the edges of the distribution. Second, as a result of physiological nystagmus, centrally located receptors are stimulated more efficiently because the rate of application of the light energy is maximized in those receptors located near the maximum slope of the distribution of light (Rushton, 1935). If we assume the primary "carrier" of contour activity to be the on-off fibers, then these fibers serving receptors centrally located will have a higher rate of firing than those serving receptors found at the periphery of the population of receptors involved. Third, because of the faster rate of application of stimulation just noted, the amount of neural accommodation in these centermost fibers will be minimized. Neural accommodation, according to Nernst (1908), is a counter-process of some kind in the neuron which cuts short the effect of a stimulus. As the rate of application of a stimulus is increased, the amount of accommodation is decreased (Monnier, 1934). Whether neural accommodation actually operates in the vertebrate retina remains to be demonstrated conclusively, but there is some evidence to suggest that it does (Granit & Skoglund, 1943). The result of the increased rate of application of the stimulus, and hence decreased accommodation, produces an additional relative increase in excitation rate for the neurons associated with the center of activity in the retina (Skoglund, 1942).

There are, in addition, possible bases for peaking the excitation through the action of retinal receptive fields (Barlow, 1953a, 1953b; Hartline, 1940; Kuffler, 1953). A receptive field of a particular ganglion cell typically is organized with antagonistic discharge patterns of the center and the periphery, that is, the center may be predominantly "on" and the periphery of the field "off," or vice versa. The intermediate area of the field will be "on–off." Thus, any displacement of a border across such a field as might be produced by slight movements of the eye will result in continuous excitation of the ganglion cell. If the field is an 'off-center" type and the gradient of stimulation passes from center to periphery then there will be bursts of excitation on leaving center, entering and leaving the intermediate area, and upon entering the peripheral part of the field. Any kind of scanning back and forth across this field obviously will maximize excitation of the ganglion cell associated with the receptive field. If to these complex events occurring at the retinal level we add the effects of reciprocal overlap occurring during projection, the distribution of excitation in the visual cortex is peaked further and we have the leptokurtic distribution as shown in Figure 1d.

The phenomenon. In our early experiments, a stimulus (A) was observed moving horizontally through distance (D), from left to right. Its velocity (V) was increased until the edges of A no longer could be seen clearly. At this velocity, if A were exposed briefly in a fixed position prior to movement (this duration we called

t_1), A was seen with sharp edges. This effect was systematically investigated for velocities from 0 to approximately 30 degrees of visual angle per second. When values of V were plotted against values of t, a linear relationship was obtained. In order to maintain the contour of a moving stimulus, each increase in V of 1 degree visual angle per second required a corresponding increase in t of approximately 25 milliseconds.

The addition of the prior stationary phase to movement definitely facilitated the perception of the edges of the moving stimulus. Stemming from a statistical treatment of apparent movement (Osgood, 1953, pp. 245–47), wherein two separate distributions of excitation in the cortex sum into one distribution whose peak is displaced in cortical space, a tentative explanation of our contour data was offered. The hypothesis was that excitation arising from the moving stimulus summed with the decaying activity around c_1 (the cortical locus of maximum activity produced by the fixed stimulus) to generate a single peaked distribution moving through cortical space toward locus c_2 (the cortical locus of maximum activity produced by the moving stimulus at the termination of movement).

THE MODEL

It is now appropriate to state some assumptions upon which our statistical concept of dynamic contour is based. Some of these assumptions follow closely those made by other statistical theorists; others are especially germane to the question of dynamic contour.

1. On-off and off receptors are the principal retinal mediators of contour perception. The excitation ensuing from the stimulation of these two types of receptors becomes the cortical representation of a border. Because of their response characteristics it is assumed that on fibers play a minor role.

2. Each of an infinite number of points along a border of a stimulus (hairline) is represented in the visual cortex by a normal or near-normal distribution of excitation (rate of firing of neural elements extended over cortical space).

3. The continuum of points along the border of the stimulus is represented in the cortex by a continuum of points of maximal excitation, the longitudinal extent of which is proportional to the length of the border.

4. The mean slope of the cortical excitation distribution, rather than the absolute amount of excitation, defines the clarity of an edge. The sharpness of the edge is determined directly by the relationship between the amount of excitation at its maximum and the amount of dispersion of excitation through cortical space. It is assumed that the mean slope of the distribution of excitation increases through time in a negatively accelerated fashion. Only when the slope equals or exceeds some critical value range is contour perceived. This implies some finite period of time as a necessary antecedent to contour perception.

The slope of the excitation distribution in the cortex at any instant in time is determined by the intensity of light, contrast, and by the state of the eye and proximal neural elements.

5. Overlapping distributions of excitation in the cortex summate.

6. The mean slope of the distribution of excitation also is influenced by velocity of movement: slope is inversely related to velocity.

7. After cessation of stimulation arising from a fixed or moving border, the mean slope of the excitation distribution decreases through time in a

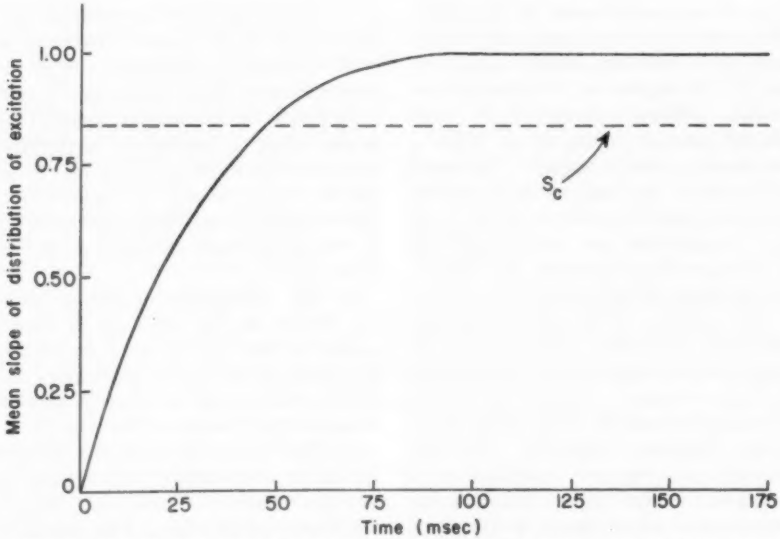


FIG. 2. Hypothetical curve depicting the increase in the mean slope of the excitation distribution through time. (S_c—critical slope—signifies the value necessary for the perception of contour. The slope of the function in the figure would change with various stimulus conditions. In this example, contour is perceived about 50 milliseconds after initiation of stimulation.)

negatively accelerated fashion. The mean slope at any given instant in time during decay is influenced by the intensity of the light: the higher the intensity, the steeper the slope.

When light stimulates the retina, the level of cortical excitation related to it grows in a negatively accelerated manner. Moreover, and of more importance, the excitation distribution representing a contour increases in mean slope through time in a negatively accelerated fashion. This increase in slope is shown in Figure 2.

This hypothetical function indicates that in approximately 50 milliseconds the slope (S) reaches a value necessary for sharp contour perception (S_c). Although this temporal value is to some extent arbitrary, it illustrates here that the appearance of sharp edges in the visual field takes a finite period of time to develop. There is no single instant when a sharp edge appears. Rather, the transition from "fuzziness" to a clearly defined edge is a continuous matter. The value of 50 milliseconds in Figure 2 represents a general threshold estimate for the appearance of a well-defined edge. Numerous factors can influence the rapidity with which S_c is attained, among which are intensity level, degree of contrast, and level of retinal adaptation. Regardless of specific conditions of observation, however, the delay in reaching S_c after the onset of stimulation is due to the temporally dispersing factors of synaptic delay, reverberation, and horizontal spread of neural activity as it proceeds from the receptors to the visual cortex.

When a border is removed from the visual field, its related excitation distribution decays with negative ac-

celeration. Furthermore, the slope of this distribution also decreases and in the same general manner (Assumption 7). Figure 3 represents a hypothetical decay function showing the decrease in mean slope of the excitation distribution as a function of time. Note that this decay function is influenced to an important degree by the duration of the stimulus and its intensity. The prolonged decline in the mean slope is due to the continued excitation as a result of the reluctance of the on-off and off fibers in the optic nerve to cease activity. The nature of this decay is based primarily upon evidence obtained by Hartline (1938) of residual activity persisting for many seconds, although gradually subsiding. Because the slope during decay exceeds S_c for some brief but finite period of time after the cessation of stimulation, an edge can be perceived during this brief interval.

The hypothetical functions depicted in Figures 2 and 3 represent the increase in mean slope as a function of time produced by an initial stationary stimulus (Figure 2) and the decrease in mean slope as a function of time produced by the cessation of the same stimulus (Figure 3). When a *moving* stimulus is presented to the eye, both the mean slope of excitation and the level of excitation increase in the same general manner as that already discussed for stationary stimuli. In Figure 4 are shown three hypothetical functions describing the mean slope, one for each of three velocities of movement. Assuming other conditions constant, they illustrate that the time required to attain S_c is determined by velocity of movement. In Assumption 6 it was stated that the mean slope of the distribution of excitation produced by a stimulus in motion was affected by the velocity of

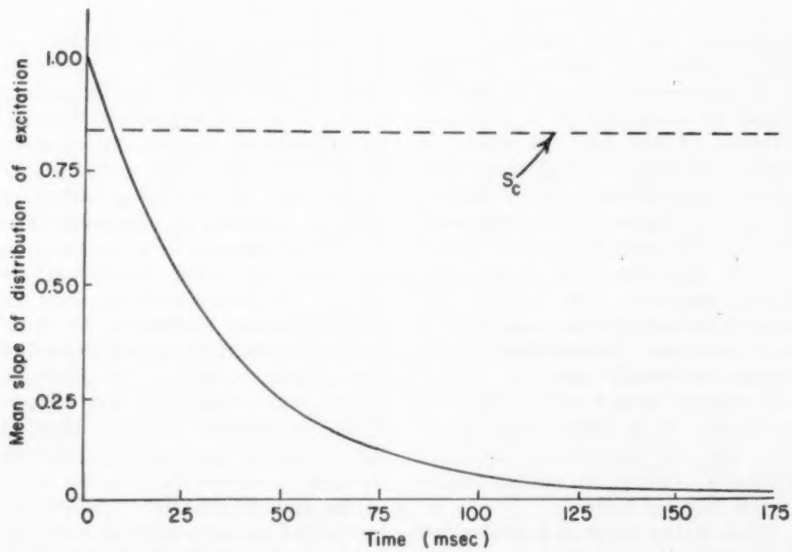


FIG. 3. Hypothetical curve of decay in mean slope of excitation distribution through time. (In this example, contour is perceived for a few milliseconds after cessation of stimulation until the slope of the excitation distribution falls below S_c.)

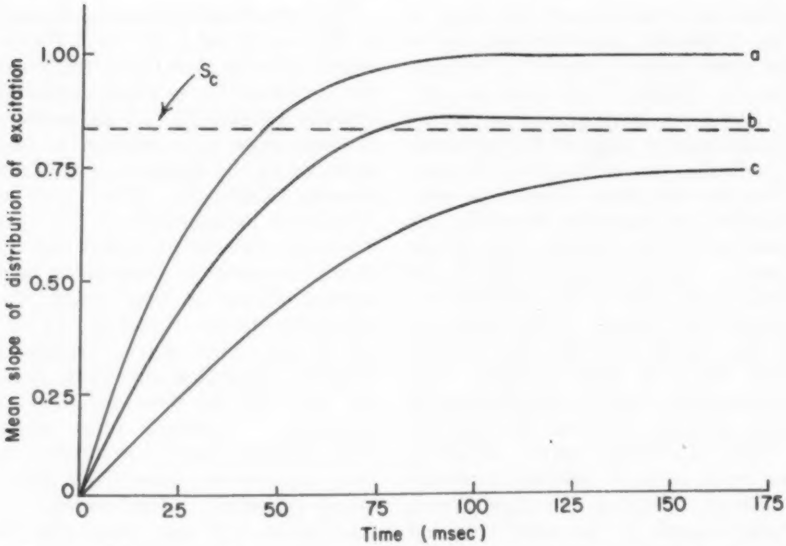


FIG. 4. Three theoretical curves showing how the increase in mean slope of the excitation distribution changes over time for different velocities of movement. (As velocity of movement increases, more time is required for the mean slope to reach S_c, the level necessary for contour perception. At high velocities S_c may not be attained—Curve c.)

movement. At low velocities ($V < 10°$/sec) the time required to attain S_c is approximately equal to that required to attain it with a stationary stimulus (Curve a). However, at higher velocities ($V > 10°$/sec) more time is required before S_c is attained (Curve b). Indeed, if V is sufficiently high, S_c will never be reached (Curve c). As illustrated with these hypothetical functions, when velocities are very high (Curve c) the mean slope of the excitation distribution never reaches the critical value, S_c. Therefore, clearly defined edges cannot be perceived. It is under these conditions that the presence of the initial stationary phase of the stimulation sequence becomes useful.

What is now required is to elucidate the manner in which the excitation arising from the stationary stimulus sums with the excitation produced by the subsequent moving stimulus so as to maintain jointly a value of S_c which is sufficient for contour perception. It is our hypothesis that the maintenance of contour during stimulus movements as a result of the presentation of the initial stationary phase is achieved by the summation of the distributions of excitation arising from the total stimulation sequence. Excitation resulting from the moving stimulus is added to the residual remaining after the withdrawal of the stationary stimulus, and their sum under certain conditions can be sufficient to increase the mean slope of the combined excitation distribution to equal or surpass S_c at a time much earlier than would otherwise have obtained had S_c been attained solely on the basis of excitation from movement.

Even under conditions of static contour perception a brief interval is re-

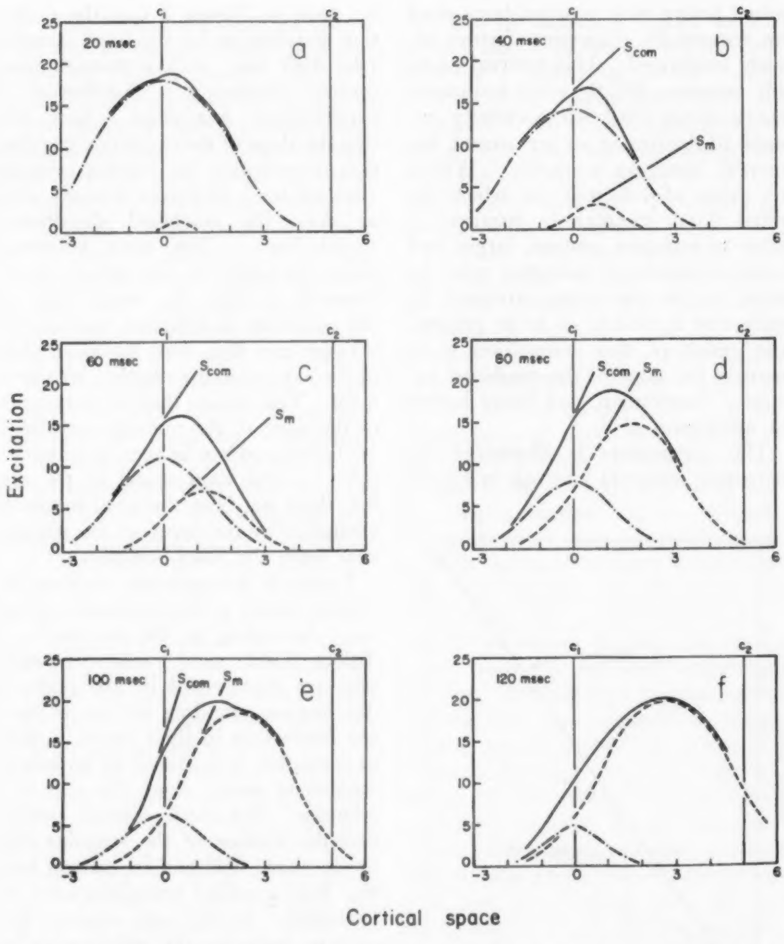


FIG. 5. Theoretical curves showing how the combined distributions of excitation for the moving contour (regular broken lines) and residual and decaying excitation from the initial fixed exposure of the stimulus (dot-dash lines) combine (solid curves) over time to produce contour perception. (As time increases from the moment of the initiation of movement to 20 milliseconds, a; 40 milliseconds, b; 60 milliseconds, c; etc., the level of excitation related to the moving border increases as it moves through cortical space. The level of excitation of the combined distribution also increases through time. Here it may be seen that the slope of the distribution representing the moving stimulus, S_m, increases through time. The slope of the combined distribution, S_{com}, increases through time at a faster rate, and thereby hastens the attainment of the critical slope, S_c, for contour resolution. The symbol C_1 represents the cortical location representing the peak of excitation from the stationary phase, and C_2 the cortical location of the peak of excitation from the moving phase at the end of movement. The abscissa represents cortical space in arbitrary units.)

quired before S_e is reached because of the temporally dispersing factors already mentioned. This interval probably increases slightly when a stimulus moves slowly; but when velocity exceeds 10 degrees or so per second, the interval increases markedly. Within the range of velocities for which the initial fixed stimulus is required in order to maintain contour, larger and larger residuals of excitation must be added to the excitation produced by movement if contour is to be present. The result of this summation is to increase the slope of the combined excitation distribution, and hence hasten the attainment of S_e.

This summation is illustrated for increasing temporal intervals in Figure 5. Note in Figure 5 that the excitation distribution for the fixed stimulus (dot-dash line) decays through time, thereby producing a distribution of progressively less slope. Note also that the slope of the excitation distribution representing the moving stimulus (dashed line) increases through time, as does the combined distribution (solid line). The most important point illustrated by the set of curves, however, is that the mean slope of the combined distribution increases at a faster rate than does the mean slope of the excitation distribution for movement. This means that the perception of the edge of the moving stimulus is facilitated and can be seen at an earlier point in time than would be the case had there not been the contribution to excitation by the decay of the stimulation from the fixed stimulus.

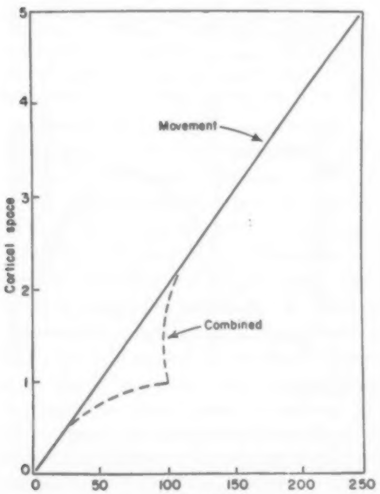


Fig. 6. When the velocity of movement of a stimulus is constant, the rate of displacement of the peak of excitation in the cortex is also constant (solid line). However, when the decaying activity from the fixed stimulus sums with the growth of activity produced by the moving stimulus, the peak of the combined distribution is not displaced through cortical space at a constant rate (dashed line).

Figure 5 demonstrates another important aspect of this summation process. According to the Marshall and Talbot (1942) view, borders of a stimulus are represented by the peaks of distributions standing for points along the borders or in their terms, a ridge of excitation is produced by an infinite number of points along the edge of a stimulus. For static contour perception the location of the stimulus edge in the visual field and its cortical location bear a rather straightforward relationship. In dynamic contour perception, however, this relationship becomes more complex. If we assume the eye to be fixed, when the velocity of movement of the stimulus object is constant, so is the velocity of the moving peak of excitation as it travels through cortical space. In dynamic contour perception based upon the summation of excitation distributions the peak of total excitation does not travel through cortical space at a constant velocity. In Figure 5d the peak of the combined distribution can be

seen to lag behind the peak produced by the moving border alone (by about one unit of cortical space). The difference in the rates of displacement for the peak of the distribution of excitation produced by the moving border and the peak of the summed distribution of excitation are presented in Figure 6. This figure illustrates that the rate of displacement of the peak of excitation for the moving stimulus alone is constant; hence the straight line. However, the rate of displacement of the peak of excitation for the combined distribution is not constant, but first negatively accelerated and then positively accelerated. It is hypothesized that this change in rate of displacement of the peak of the distribution of the summed excitations is responsible for a peculiar effect reported by our subjects; namely, that the velocity of the moving edge did not appear to be constant. This account of perceived variable velocity with constant velocity of movement we believe to be one important piece of supporting evidence for the theory presented here.

To summarize: For low velocities of movement S_c is attained with sufficient speed to allow virtually instantaneous contour perception. With high velocities S_c is never attained during the time interval required to complete the stimulus movement sequence. For the intermediate range of velocities (8–30°/sec), S_c is attained rapidly enough to permit contour perception only if the decaying neural activity produced by the fixed stimulus sums with the growth of neural activity produced by the moving stimulus.

SUPPORTING EVIDENCE

Lack of techniques for specific measures of the integrative activity of the visual system makes difficult any direct test of statistical theory as expressed here. However, indirect evidence of a supporting nature can be obtained from various perceptual experiments. We have mentioned previously the evidence concerning the perceived variable velocity of a constant velocity stimulus. To other supporting evidence we now turn.

Contrast. We have said that contour perception cannot occur unless the slope of the cortical distribution representing contour reaches some minimal value (S_c). One factor which should have a direct and important bearing on the activity of on-off fibers, and hence the peaking of the related cortical distribution of excitation, is the intensity difference representing a border. It was hypothesized, therefore, that under conditions of high contrast between a stimulus object and its surround, the perception of a moving edge could occur at velocities of movement higher than those which would be expected under low contrast. Because high contrast would decrease the time required to obtain S_c, the importance of the summative effects of decaying neural activity following the presentation of the initial stationary stimulus would be minimized. That is, a stimulus of high contrast at a given velocity should require less contribution from the decaying effects of the initial stationary stimulus than a stimulus of the same velocity and lower contrast. The major contribution of the initial stationary stimulus is to provide, after its cessation, neural excitation which may sum with that produced by the moving stimulus. This summation effect simply hastens the attainment of S_c during movement (see Figure 5).

In terms of our operations the result of an increase in contrast would be to allow the perception of contour of the moving stimulus at a higher

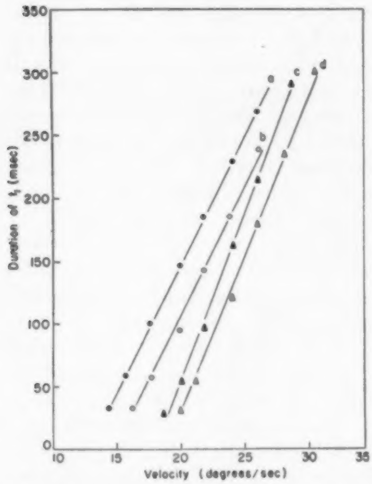


FIG. 7. The effects of contrast and illumination level on contour perception of moving stimuli. (Data of five subjects. Brightness values in footlamberts of the background and the stimulus: d, 4.2–1.2; c, 4.2–1.5; b, .036–.007; a, .036–.015. The stimulus was always darker than the background. Stimulus size equaled 0.5 degrees, and the extent of movement was 5.0 degrees of visual angle. Each curve in the figure specifies how long the stimulus had to appear in a fixed position prior to movement, t_1, in order for contour to be perceived during movement of a particular velocity.)

velocity than under conditions of low contrast, even though the duration of the fixed stimulus remained constant. Conversely, for a given velocity of movement, a reduced exposure interval of the initial fixed stimulus is adequate to maintain contour perception during movement whenever contrast is increased.

Data on the effects of contrast are presented in Figure 7. For each of five subjects, velocity of movement was increased from 8 degrees per second in 2-degree steps until contour was lost (about 13°/sec). Thereafter, the value of t_1 required to see contour was determined for each increase in V. In Figure 7 mean values of t_1 are plotted against V for each of two contrast conditions at each of two levels of illumination. These data show that increased contrast selectively displaces the time-velocity contour function in the expected direction: higher contrast enables the subject to perceive the edge of the moving stimulus at higher velocities with a constant value of t_1. Under low illumination the increase in contrast shifted the function from a to b, and under high illumination, the increase in contrast shifted the function from c to d.

Intensity level. The data plotted in Figure 7 also show an effect one would predict from statistical theory; namely, that contour perception of stimuli in motion is influenced by intensity. In Assumption 7 we stated that the mean slope of the decaying distribution of excitation became steeper as stimulus intensity increased. Furthermore, the rise in the mean slope of the distribution of excitation from the moving stimulus with increased intensity will be faster and will hasten the attainment of S_e. Accordingly, the importance of the summative effects of the initial stationary stimulus would be reduced.

As seen in Figure 7, the result of increasing intensity shifted the time-velocity contour function in the expected direction; under low contrast, increased intensity shifted the function from a to c, and under high contrast an increase in intensity shifted the function from b to d.

Maintaining a constant contrast between the stimulus and its surround while increasing intensity should and does facilitate contour perception, but only up to a point. If adjacent areas of a visual pattern differing in intensity have their intensities increased proportionately, a level will be reached where the necessary conditions for sta-

tistical peaking will no longer be possible. Thus if the areas defining a contour are both sufficiently intense, on-off receptors would not respond and the presumed normal distribution of excitation would degenerate.

Spatial displacement. Our account of how decaying neural activity from fixed stimulation combines with that produced by a moving stimulus requires particular spatial and temporal relationships of these two parts of the stimulation sequence. Any change which destroys the spatial and temporal contiguity of the two phases should have an adverse effect upon contour formation.

We predicted that the introduction of a spatial separation in the visual field between the position of the fixed stimulus and the initial position of the moving stimulus would require a higher value of t_1 to allow contour perception during movement than would be required in a control condition without spatial separation. The rationale for the separation in the visual field of the two phases of the stimulation sequence was to separate the assumed distributions of excitation produced by them in the visual cortex. It was necessary to provide the subject with a minute fixation point located at the place where movement began, because without it the subject probably would have tended to view the entire stimulus sequence with the same area of the retina. Following the cessation of the stationary stimulus, the moving stimulus appeared at the point of fixation and moved from left to right.

If the neural elements involved in the excitation produced by a stationary stimulus comprised a totally different population of elements from those involved in excitation from a moving stimulus, then no summation could take place. On the other hand, maximum summation would occur when identical elements comprised both populations. It was predicted that the greater the spatial displacement (the less the summation), the more adverse would be the effect upon contour formation.

Data obtained confirmed the spatial displacement hypothesis. Figure 8 presents the time-velocity functions obtained under several experimental conditions. These data are from eight subjects, and the procedure employed in obtaining the data was the same as that indicated previously. (Differences in the origin of the time-velocity functions which exist from one experiment to another are assumed to be due to individual differences, and differences in illumination and contrast.) The control condition consisted of having the stationary stimulus appear at the same place in the visual field from which movement began. Displacements were measured in degrees of visual angle at 1 and 2 degrees right, and 1 and 2 degrees left. Displacements of equal extent to the right and left produced equivalent shifts in the functions and so data were combined across this variable.

Temporal displacement. If there were a time lapse between the cessation of the initial stationary stimulus and the onset of the moving one, then subsequent summation or neural activity should be reduced. This reduction in summation would increase the time required to obtain S_c and would minimize the influence of the stationary stimulus upon the perception of contour of the moving phase.

Judgments of contour relating to this issue were obtained for eight subjects under each of three conditions. In the control condition the two phases of the stimulus sequence were temporally contiguous, where in the other two conditions temporal delays of 40

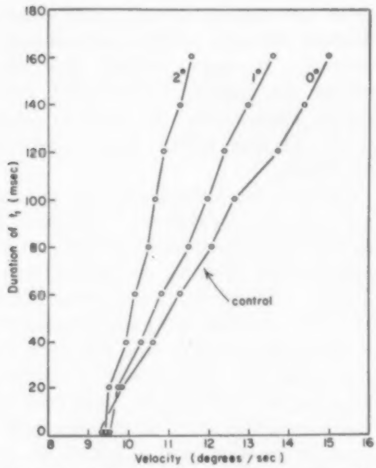


FIG. 8. The effects of spatial separation of the two phases of stimulation on dynamic contour. (Combined data of eight subjects. Stimulus brightness, 2.0 footlamberts; brightness of background, 70.0 footlamberts. Stimulus size, 0.5 degree. Extent of movement, 5 degrees of visual angle. Increasing the spatial separation between fixed phase of stimulation and moving phase required a greater exposure duration of the fixed phase in order to retain contour perception for a particular velocity of movement. Displacements to the right and left produced equivalent effects. Therefore, data were combined across this parameter.)

and 80 millseconds were employed. We predicted that the temporal delays would shift the time-velocity functions to the left indicating, that for a particular velocity, longer exposures of the stationary stimulus would be required to produce contour formation. The amount of neural activity available to sum with that produced by the moving stimulus is determined not only by temporal displacement of the two phases of the stimulus sequence, but by the duration of the stationary stimulus as well. Accordingly, we hoped to obtain data suggesting that what is lost to the summative process as a result of temporal delay can be recovered by increasing the value of t_1.

Data relevant to this problem are presented in Figure 9. The shift in the contour function as a result of temporal displacement is apparent. The greater the delay, the larger the shift to the left from the control condition. Presumably, a delay longer than 80 milliseconds might produce a time-velocity function asymptotic to the ordinate, suggesting that beyond a certain duration of temporal displacement, increased durations of t_1 would cease to influence the perception of contour. Increased values of t_1 *can* compensate for temporal displacement provided that the latter is not too long. For example, at a velocity of 18 degrees per second the duration of the sta-

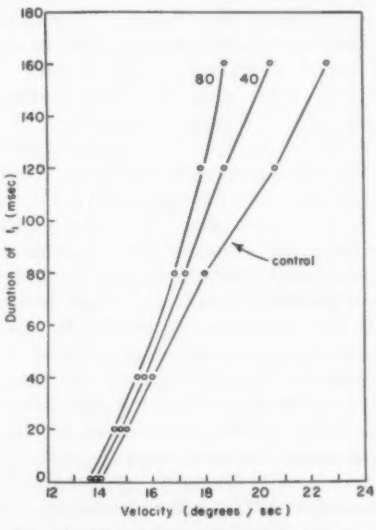


FIG. 9. The effects of temporal separation of the two phases of stimulation on dynamic contour. (Combined data of eight subjects. Brightness of stimulus, 2.0 footlamberts; brightness of background, 70.0 footlamberts. Stimulus size, 0.5 degree of visual angle. Extent of movement, 5 degrees. For a given velocity of movement, the exposure duration of the fixed phase of stimulation must be progressively increased as temporal separation increases in order to retain contour.)

tionary stimulus necessary for contour perception during movement was 80 milliseconds for the control condition (no temporal displacement). However, with temporal displacements of 40 to 80 milliseconds, the durations of the fixed phase had to be increased to approximately 100 and 120 milliseconds, respectively.

DISCUSSION

The conception of contour perception which we have offered, while congruent with certain facts concerning the visual system and with particular experimental results, has not directed itself to various considerations of a complicating and relevant nature. These considerations include the question of eye movement, the locus of the summation process, the question of neural inhibition, and some general questions about neurophysiological mechanisms.

Eye movements. The perception of edges of moving stimuli obviously raises the problem of how ocular tracking is related to the formation of contours. Surprisingly little empirical knowledge is available regarding the nature of ocular tracking, although one of the authors is currently investigating this problem. Having available a photoelectric system of measuring eye movements which is coupled to a system for generating moving targets, precise measures can be made of ocular tracking and these then related to perceptual judgments of targets whose velocity, displacement, and other attributes can be varied (Smith & Warter, 1959, 1960). Investigations to date indicate quite clearly that for the border of a moving target to be seen clearly a high degree of accuracy in visual tracking is *not* required. In fact, it appears that it is only for the very slow velocities (constant) of movement (about 10°/sec or less) that the eye tracks with reasonable accuracy. The upper limit of 10 degrees per second, it is important to note, is approximately the velocity of movement at which the borders of the moving stimulus begin to be seen as blurred when the duration of the stimulus in a fixed position before movement (t_1) is zero. It will be recalled that the most important single fact reported in our contour experiments is that by increasing the interval of exposure of the stimulus in a fixed position prior to movement we can systematically increase the upper limit of velocity at which the edge of the stimulus still can be seen clearly. A basic question raised by our experiments concerns how an increase in t_1 affects visual tracking. A provisional answer to this question based upon eye movement data to date is that increasing the magnitude of t_1 improves visual tracking generally but not to a level of accuracy characteristic of tracking at low velocities. About all we can conclude at this point is that the facilitation of dynamic contour perception by exposure of the target in a fixed position prior to movement is in some manner directly related to the accuracy of visual tracking. We can be sure that in our contour experiments the displacement of the stimulus at a constant velocity did not produce on the retina a displaced locus of maximum excitation which also was displaced at a constant velocity. We must assume that the cortical representation of a contour must contain components of acceleration and deceleration not found in the original stimulus. This, of course, is due to irregularities in the pursuit movement of the eyes under conditions wherein the moving stimulus alone is presented, and it is due to inaccurate tracking and the summation process under conditions in which both

phases of the stimulation sequence are presented (see Figure 6).

Based upon eye movement data we can offer a suggestion concerning the possible relation between dynamic contour perception and correlated visual tracking movements. If one examines the magnitudes of lag and lead errors of the eye as the velocity of a stimulus increases, it is obvious that the magnitude of these errors increases as velocity increases, and at any given velocity it decreases at t_1 is increased. It is our hypothesis that as long as the lag or lead error does not exceed a few degree of angle (3–4 degrees at most), contour perception can be maintained. If this hypothesis is supported by the experiments now in progress, it will converge nicely with the facts about the maximum angular separation over which optional apparent movement can be obtained (De Silva, 1929), and the maximum distance over which one contour can exert an inhibitory influence over another (Fry & Bartley, 1935). Moreover, it is very interesting to note, and perhaps very important that the figure of 3–4 degrees angular magnitude is just about the visual field size represented by the measured size of receptive fields as revealed by Kuffler (1953).

The data obtained so far on the problem of eye movements tend to complicate the model, but do not vitiate it. Since the data on eye movements are at present incomplete we have assumed in the description of the summation process (see Figure 5) that the eye lags behind the target in the hypothetical situation depicted. If our eye movement data provide us with data contrary to this assumption, then the effect would be to augment the summation process by keeping the growth of neural excitation produced by the moving stimulus nearer the population of neural elements undergoing decay following the cessation of the stationary stimulus. Again, the important point is that the general statement of the model can accommodate the complicating factor of eye movements.

Locus of summation. Since the first statement of statistical theory by Marshall and Talbot (1942), they and subsequent statistical theorists have placed the locus of contour resolution in the striate cortex. The primary role of the projection system has been seen as contributing to the peaking of the distribution of neural excitation. The locus of summation of different distributions of excitation, however, has been the visual cortex alone. The question may be raised, however, as to the possibility of summation at lower levels. In an effort to ascertain the likelihood of this possibility we executed an experiment in which the stationary stimulus was presented to one eye whereas the moving stimulus was presented to the contralateral eye (Gulick, 1960). The results of this experiment suggest that for maximum contour facilitation it is necessary to present the stimulation sequence in such a way as to allow neural interaction and/or summation at the retinal level as well as at the cortical level. The implications of these data for the model involve a possible extension of the present view of neural summation as it is conceived in current statistical theory.

The interaction (and subsequent neural summation) of two temporally and spatially separate visual events which is central to our position raises in addition certain questions which require further investigation. We have presented evidence concerning the effects of separating in space and time the initial exposure of the stimulus in a fixed position and the movement component. It is necessary to ask

what effects upon the perception of edge of the moving stimulus would be observed if the fixed and moving components of presentation differed in terms of certain attributes which have yet to be studied. For example, suppose that the stationary stimulus prior to movement were exposed out of focus, to be followed by the moving phase with the stimulus in focus. What then would be the change in our typical relation between velocity and t_1 as far as perceiving the moving edge is concerned? Similar important questions also need to be studied once we have completed the instrumentation necessary to present stimuli in this manner.

Neural inhibition and facilitation. As a final point of discussion we realize, as must the reader, that our conception of contour and its expression in neurological and physiological terms is unbalanced. The theory rests principally upon the concept of neurol summation and says little, if anything, about related concepts of inhibition and facilitation. In view of the inhibitory effects that one contour can have upon an adjacent one when both are fixed in position (Fry & Bartley, 1935), we are acutely aware that similar effects may be present when a particular contour is displaced relative to its original position. We think it reasonable, for example, to assume that an inhibitory mechanism can account for the "damping" of residual retinal activity correlated with the "trailing" edge of any moving stimulus. In any event we assume that future and more sophisticated expressions of statistical theory will incorporate modes of neural interaction in addition to summation at the cortex.

Regarding possible future expressions of statistical theory of contour we can indicate briefly some of the things we believe ought to be done. Additional perceptual studies are needed in which systematic investigations of the effects of differences in the fixed and moving components of the stimulus presentation are made. Moreover, we need to know more about the relationships between tracking motions of the eye and various conditions of stimulus movement. Finally, and most important, direct evidence bearing upon the neurological basis of perception is required if we are to continue to use with confidence the statistico-neurological assumption upon which the theory rests. While congruent with some neurological evidence, statistical theory of contour perception is far from being firmly based upon unambiguous neurological and physiological data. And, of course, the theory cannot find its real tests in perceptual data alone. We must continue to search for direct measures of the integrated activity of the visual system itself in relation to the data of perception. By so proceeding we will determine if statistical theory will prove itself as a fruitful means of conceptualizing basic problems in perception.

REFERENCES

Barlow, H. B. Action potentials from the frog's retina. *J. Physiol.*, 1953, **119**, 58–68. (a)

Barlow, H. B. Summation and inhibition in the frog's retina. *J. Physiol.*, 1953, **119**, 69–88. (b)

Bartley, S. H. *Vision: A study of its basis.* New York: Van Nostrand, 1941.

Day, R. H. Application of the statistical theory to form perception. *Psychol. Rev.*, 1956, **63**, 139–148.

Day, R. H. The physiological basis of form perception in the peripheral retina. *Psychol. Rev.*, 1957, **64**, 38–48.

De Silva, H. R. An analysis of the visual perception of movement. *Brit. J. Psychol.*, 1929, **19**, 268–305.

Deutsch, J. A. The statistical theory of figural after-effects and acuity. *Quart. J. exp. Psychol.*, 1956, **47**, 208–215.

Fry, G. A. *Blur of the retinal image.* Columbus, Ohio: Ohio State Press, 1955.

Fry, G. A. & Bartley, S. H. The effect of one border in the visual field upon the threshold of another. *Amer. J. Physiol.,* 1935, **112**, 414–421.

Granit, R., & Skoglund, C. R. Accommodation and autorhythmic mechanisms in single sensory fibers. *J. Neurophysiol.,* 1943, **6**, 337–348.

Gulick, W. L. Monocular contour perception under the influence of prior stimulation of the contralateral eye. *Psychol. Rec.,* 1960, **10**, 123–130.

Hartline, H. K. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Amer. J. Physiol.,* 1938, **121**, 400–415.

Hartline, H. K. The nerve messages in the fibers of the visual pathway. *J. Opt. Soc. Amer.,* 1940, **30**, 239–247.

Köhler, W., & Wallace, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. Phil. Soc.,* 1944, **88**, 269–357.

Kuffler, S. W. Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.,* 1953, **16**, 37–68.

Marshall, W. H., & Talbot, S. A. Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In H. Klüver (Ed.), Visual mechanisms. *Biol. Symp.,* 1942, **7**, 117–164.

Monnier, A. M. *L'excitation électrique des tissues: Essai d'interpretation physique.* Paris: Hermann, 1934.

Nernst, W. Zur Theorie des elektrischen Reizes. *Pflüg. Arch. ges. Physiol.,* 1908, **122**, 275.

Osgood, C. E. *Method and theory in experimental psychology.* New York: Oxford Univer. Press, 1953.

Osgood, C. E., & Heyer, A. W., Jr. A new interpretation of figural after-effects. *Psychol. Rev.,* 1952, **59**, 98–118.

Rushton, W. A. H. The time factor in electrical excitation. *Biol. Rev.,* 1935, **10**, 1.

Skoglund, C. R. The response to linearly increasing currents in mammalian motor and sensory nerves. *Acta physiol. Scand.,* 1942, **4**, Suppl. No. 12.

Smith, K. The statistical theory of the figural after-effect. *Psychol. Rev.,* 1952, **59**, 401–402.

Smith, W. M., & Gulick, W. L. Visual contour and movement perception. *Science,* 1956, **124**, 316–317.

Smith, W. M., & Gulick, W. L. Dynamic contour perception. *J. exp. Psychol.,* 1957, **53**, 145–152.

Smith, W. M., & Warter, P. J., Jr. Photoelectric technique for measuring eye-movements. *Science,* 1959, **130**, 1248–1249.

Smith, W. M., & Warter, P. J., Jr. Eye movement and stimulus movement: New photoelectric electromechanical system for recording and measuring tracking motions of the eye. *J. Opt. Soc. Amer.,* 1960, **50**, 245–250.

(Received August 23, 1960)

Psychological Review
1962, Vol. 69, No. 2, 109–135

SUBJECTIVE PROBABILITIES INFERRED FROM DECISIONS[1]

WARD EDWARDS

Engineering Psychology Laboratory, University of Michigan

Theories about how people make decisions in risky or uncertain situations have come to focus on two concepts: utility, or subjective value, and subjective probability. The status of the utility concept is fairly clear (for reviews, see Edwards, 1954c, 1961). But the status of the subjective probability concept is quite confused, both theoretically and in relation to experiments. Too often the concept of subjective probability has been introduced casually into a decision theory which focuses on the concept of utility. In fact, the only thorough treatment of a concept of subjective probability (quite different from the one to be discussed in this paper) is L. J. Savage's (1954) brilliant book. The notion of subjective probability would be necessary in psychology even if it played no role in decision theories, since men often make judgments about probabilities based on displays of random or nonrandom events. But this paper is mostly concerned with subjective probabilities inferred from decisions via some kind of value or utility maximization model of the decision-making process.

Psychologists trained in psychophysics tend to think that subjective probability is related to objective probability in more or less the same way that the subjective loudness of a tone is related to its objective intensity. The purpose of this paper is to examine the content, merits, and limitations of such an approach. The discussion will focus on two closely related matters. The first is the idea of a set of functions relating subjective to objective probability. The second is whether or not the subjective probabilities of a set of mutually exclusive events, one of which must happen, should add up to one.

The paper begins by defining two classes of decision theories. After some preliminary discussion of utility and subjective probability functions, it next considers the class of theories which result when subjective probabilities are assumed to add up to one. This class turns out to have some serious difficulties. A brief review of experimental evidence provides empirical reasons for avoiding these difficulties by rejecting additivity. Then the paper examines models which do not require subjective probabilities to add up to one. Such models require a reformulation of the concept of utility; utility scales must have true zero points. Finally the paper derives two experimentally testable

[1] The work reported here was supported in part by the Operator Laboratory, Air Force Personnel and Training Research Center, and in part by Project MICHIGAN under Department of the Army Contract DA-36-039-SC-78801, administered by the U. S. Army Signal Corps. Reproduction for any purpose of the U. S. Government is permitted. Credit amounting to coauthorship should go to L. J. Savage, W. M. Kincaid, and George Minty, who did the mathematical work on which this paper is based; their individual contributions are acknowledged where they occur. I am also deeply grateful for advice and criticism to them and to many others, especially G. M. Becker, C. H. Coombs, R. D. Luce, Jacob Marschak, Anatol Rapoport, M. Toda, E. L. Walker, and K. V. Wilson. An earlier version of this paper (Edwards, 1959, 1960) is closely related.

properties which any plausible utility-subjective probability model, additive or otherwise, must have, and presents a model which I find attractive.

MAXIMIZATION AND UTILITY

First, some definitions and assumptions. This paper is concerned with decisions of the following kind. A decision maker can choose one and only one of two or more courses of action. Associated with each course of action is a finite set of possible outcomes, one and only one of which will happen if that course of action is chosen. Each possible outcome can be described by a number, called a utility, defined at least up to a linear transformation (i.e., measured at least on an interval scale); this number fully represents the desirability of that outcome to the decision maker. If more than one possible outcome is associated with a given course of action, then a well defined external event, over which the decision maker has no control, determines which outcome will occur if that course of action is chosen. Most nonsocial decisions fit this description.

The crucial assumption of the class of theories considered here is that people behave as if they were attempting to maximize some quantity—perhaps not skillfully, perhaps not consistently, but still attempting to maximize. Maximize what? The following equation defines the concept of subjectively expected utility (SEU):

$$\mathrm{SEU} = \sum_i \psi_i u_i \qquad [1]$$

It refers to a course of action which has a number of possible outcomes. The i^{th} possible outcome has an objective value to which corresponds a subjective value or utility u_i. Outcome i will occur if a given external event occurs; that event may have an objective probability, or the objective probabilities of one or more of the relevant set of external events may be undefined. In either case, outcome i has a subjective probability ψ_i. The theories with which this paper is concerned simply assert that subjects (*S*s) behave as though they choose, from among the courses of action open to them, the one which has the largest SEU. For more extended discussion of the origin of these models, see Edwards (1954c, 1955). This paper will treat them as models of what men do (descriptive models) rather than of what they should do (normative models).

Equation 1 introduces formally the central concept of this paper: subjective probability. What is a subjective probability? It is a number between zero and one which describes a person's assessment of the likeliness of an event. Exploration of several more explicit definitions follows.

This paper assumes that objective probabilities exist, are known for some events, and are distinct from subjective probabilities, though subjective probabilities are related to objective ones. Some theorists like Savage (1954) have asserted that there is no such thing as objective probability; all probabilities are subjective. This is no place to argue whether or not objective probabilities can be meaningfully defined. The argument is irrelevant. Some events, like the toss of a coin or the roll of a die, have well established "conventional" probabilities. Whether these probabilities are objective or not, they can meaningfully be compared with the subjective probabilities revealed when people make decisions which involve these events.

Of course many events exist which

are neither certain nor impossible but for which no objective or conventional probability can be defined. Example: Within the next 24 hours, you will drink a glass of wine. Such events are called *uncertain*, while events for which objective probabilities can be defined are called *risky*. This paper will not consider uncertain events. The same decision theory ought to apply to decisions whose outcomes are contingent on uncertain events as to decisions whose outcomes are contingent on risky events, and the latter are more helpful in examining the nature of that theory.

The notion of a set of functions relating subjective to objective probability is vital to this paper. But the theorems proved below are relevant also to theories which deny or ignore the existence of such a relationship. They imply for such theories the existence of severe restrictions on permissible relationships between subjective probabilities and classes of risky events (classes defined by objective probability values).

From here on, some abbreviations are needed. The word "maximization" will be omitted from phrases like "the SEU maximization model." The concept of objective probability will be symbolized OP in text; SP will stand for subjective probability. The concept of a set of events of which some one must happen and not more than one can happen—that is, the concept of an Exclusive, Exhaustive set of Events—will be symbolized EEE in text. Two classes of SEU models exist. Those in which the SPs of an EEE must add up to a constant will from here on be called ASEU (additive subjectively expected utility maximization) models. Those in which the SPs of an EEE do not need to add up to anything in particular (though of course the OPs of an EEE must always add up to one) will be called NASEU (nonadditive subjectively expected utility maximization) models.

The subject of transformations will come up. Psychologists speak of three classes of scales, defined by three transformation groups, which are relevant here. An interval scale permits the linear transformation $x' = ax + b$, $a > 0$. A ratio scale permits the scale transformation $x' = ax$, $a > 0$. An identity scale permits only the identity transformation $x' = x$. This paper will use the psychological rather than the mathematical names for these classes of scales.

CLASSES OF SEU MODELS

This section will identify two main classes of SEU models, ASEU and NASEU, and will then examine some extraordinary restrictions on the utility functions which must enter into each. The difference between ASEU and NASEU models is defined by

Theorem 1: If only interval, ratio, and identity scales are considered, in any SEU model in which the SPs of an EEE must add to a fixed constant (ASEU model) utility must be measurable at least on an interval scale and SP must be measurable at least on a ratio scale. In any SEU model in which no restrictions are put on the sum of the SPs of an EEE (NASEU model), utility and SP must both be measurable at least on ratio scales.

Consider the NASEU models first. Since the transformations under consideration are successive restrictions of the general linear transformation, consider the least restrictive possibility first. In that case, the transformation $u' = au + b$, $a > 0$ on utility and $\psi' = c\psi + d$, $c > 0$ on SP are permissible. Consider two courses of action, A and B, which have equal SEUs. (Continuity assumptions not

spelled out here are necessary to guarantee that such events exist and that they do not have certain pathological properties. Axioms 1 and 2 of the mathematical Appendix, plus the assumption that there are at least three possible outcomes with different utilities, would be more than enough.) Permissible transformations must not change equal SEUs, which imply indifference between A and B, to unequal SEUs, which imply that one is preferred to the other. Consequently:

$$\sum_i \psi(A_i)u(A_i) = \sum_j \psi(B_j)u(B_j) \leftrightarrow$$

$$\sum_i \psi'(A_i)u'(A_i) = \sum_j \psi'(B_j)u'(B_j) \quad [2]$$

A_i = i^{th} possible outcome of course of action A, $i = 1, 2, \cdots, I$.

B_j = j^{th} possible outcome of course of action B, $j = 1, 2, \cdots, J$.

Substituting the transformations under consideration into the second half of Equation 2 produces:

$$\sum_i [(c\psi[A_i] + d)(au[A_i] + b)] = \sum_j [(c\psi[B_j] + d)(au[B_j] + b)]$$

Multiplying and collecting terms:

$$ac[\sum_i \psi(A_i)u(A_i) - \sum_j \psi(B_j)u(B_j)] + bc[\sum_i \psi(A_i) - \sum_j \psi(B_j)] + ad[\sum_i u(A_i) - \sum_j u(B_j)] + bd[I - J] = 0 \quad [3]$$

By hypothesis:

$$\sum_i \psi(A_i)u(A_i) - \sum_j \psi(B_j)u(B_j) = 0$$

and so the first term of Equation 3 vanishes. Since both b and d may assume any arbitrarily chosen values, the only way to guarantee that the sum of the other three terms will be zero for any values of b and d is to require that the quantity inside the brackets in each term be zero. But the quantity inside the brackets in the second term (and in the third also) is explicitly not required to be zero. Consequently this set of transformations cannot be permissible. Since both a and c are required to be greater than zero, the only way to guarantee that Equation 3 will be correct is to require that both b and d be zero. That, of course, means that both utility and SP must be measured on ratio scales. Thus the second half of Theorem 1 is established.

Now consider the first half of Theorem 1, which is concerned with ASEU models. In such models, the first and second terms of Equation 3 are zero by hypothesis. Furthermore, it is clear that in such models SPs may not have an arbitrary origin, since they are required to add up to a fixed constant, and consequently d must be zero. If d is zero, the other two terms vanish, and no restrictions on b are necessary. This establishes the first half of Theorem 1.

Possible Relations between Utilities and Objective Values

Luce (1959) has recently pointed out some surprising and severe restrictions on possible relationships between objective and subjective continua of any sort, in a paper of extraordinary depth and importance. He asserts two principles to which any function relating an independent to a dependent variable within a theory should conform. First:

> Admissible transformations of one or more of the independent variables shall lead, via the substantive theory, only to admissible transformations of the dependent variables. (p. 85.) (Note that a form of this principle was used to prove Theorem 1.)

Secondly:

> except for the numerical values of parameters that reflect the effect on the dependent variables of admissible transformations of the independent variables, the mathematical structure of the substantive theory shall be independent of admissible transformations of the independent variables. (p. 85.)

The first condition means, for example, that a theory in which the dependent variable is measured on a ratio scale may *not* require that when the arbitrary origin of an interval scale independent variable is changed, the origin for the dependent variable must change also. The second, more subtle, condition means that the mathematical form of the function relating independent to dependent variables may not be changed from, for example, a power function to a simple sum of variables by permissible transformations on the independent variables. Both of these principles of theory construction seem plausible and even necessary. What do they imply for utility and SP?

First, consider interval scale utility as dependent variable and ratio scale money (or quantity of any other continuously variable good) as independent variable. The relation between these two variables, if it is to conform to the two requirements of the previous paragraph, must, according to Luce's theorems, have the form $u(\$_i) = a \log \$_i + b$, where a and b are arbitrary constants and $\$_i$ is any amount of money. If, as is necessary in NASEU models, utility is measured on a ratio scale, the relation between it and money must, according to Luce's theorems, be $u(\$_i) = a\$_i^b$ (again, a and b are arbitrary constants).

These restrictions on the possible forms of utility functions are astonishingly severe. For example, they imply that if for interval scale utilities there are any points on the function relating utility of money to its objective value at which that function increases (that is, if a is positive), and if the value of b is finite, which it must be if the equation is to make sense, then the utility corresponding to zero dollars (or some fixed positive or negative amount, since the origin may be displaced by a fixed amount) is minus infinity. The utility of negative amounts of money (or amounts less than the fixed amount) is undefined. And the form of the utility function is necessarily negatively accelerated; that is, it is impossible to find a person who attaches as much as twice as much utility to $\$2x$ as he does to $\$x$. In short, interval scale utilities associated with ratio scale money values require ridiculous utility functions.

The relation between ratio scale utilities and ratio scale money values is much less ridiculous. Its most important restrictive characteristic is that it may have only one inflection point. That means that if the function is ever positively accelerated, it is always positively accelerated; if it is ever negatively accelerated, it is always negatively accelerated. Even this characteristic is pretty restrictive. But it is not utterly ridiculous, as are the properties of interval scale utilities discussed in the previous paragraph.

These restrictions are so surprising that one is naturally tempted to try to wiggle out of them by arguing against the assumptions which lead to them. Luce (1959) has extensively discussed these assumptions and their possible limitations. This paper will not cover the same ground again; instead, I will simply record my opinion that at least for interval scales Luce's restrictions are inevitable and we must therefore learn to live with them.

Do these principles of theory con-

struction place any restrictions on possible relations between SP and OP? Probably not. We usually consider OP to be measured on an identity scale; if it is, Luce's arguments place no special restrictions on the mathematical form of possible relationships between it and anything else.

GENERAL CHARACTERISTICS OF PSYCHOPHYSICAL DECISION MODELS

Before getting down to specific issues concernings SPs, I must set the stage by discussing some general issues and some characteristics which all ASEU or NASEU models must or should have.

Identifying Characteristics

To predict choices from an SEU maximization model of any kind, you must know the utility and the SP of each possible outcome of each available course of action. You can then substitute these values into Equation 1, perform some arithmetic, and predict that the course of action with the highest SEU will be chosen. Such predictions will be wrong on occasion. This fact can be included in the model (at a heavy penalty in additional mathematical complexity) by making it stochastic, or it can be excluded from the model and left to a general theory of errors, as has been customary in most of psychophysics. This paper does the latter, but it seems likely that many of the conclusions established here for deterministic models will apply to stochastic models of the same type as well.

In order to use an SEU model, then, you need to know utilities and SPs independently of the decision you are trying to predict. How can you find them? That is a very difficult, much debated question. For the purposes of this paper, it is enough to assume that some psychophysical method is available which gives accurate utility and SP measures for any object or event of interest, and that you have available such information about anyone whose decisions you are predicting.

In what form can such information be available? Consider first information about utilities. If money is the valuable object in question, presumably a table could be prepared showing the amount of utility an *S* attaches to a gain of one dollar, of two dollars, of three dollars, and so on. (My reasons for saying "a gain of one dollar" rather than simply the amount one dollar or a total fortune of $5001 will become clearer later; the differences among these formulations are irrelevant at this point.) But the information in that table would have to be based on experiments in which the *S* received or lost particular dollar bills, with particular serial numbers. How do we know that he will respond similarly to different dollar bills, or perhaps even to nickels and dimes instead? Of course we don't. We simply assume that (within reasonable limits) the form in which *S* receives the increment of one dollar is irrelevant; that is, we abstract the property "United States currency with an objective value of one dollar" from a variety of other properties of the stimuli used in the utility measurement experiment, and assume that it is the one on which the results most significantly depend.

In the case of utility, the point made above is so familiar as to be trivial. But what about the case of SP? On what properties of events do SPs most significantly depend? The most obvious characteristic of events which should turn out to be highly correlated with SP is, of course, OP (remember that this paper is re-

stricted to classes of events for which OPs are well and independently defined). Unfortunately, unpublished data make it perfectly clear that if two bets are identical in payoff, OP, and all other characteristics except the means by which the OP is displayed, *S* may markedly prefer one to the other—so much so that he is willing to pay for the privilege of using his preferred display, even though he loses in expected value (hereafter abbreviated EV) by doing so. Within the framework of the theories discussed in this paper, this finding can be explained only by assuming that several different SPs may correspond to a given OP, depending on the nature of the OP display. In short, any attempt to state SP as a function of OP alone is doomed to failure.

What then is SP a function of? Mathematical decision theorists, aware of logical difficulties which result from thinking of SP as a function of OP, often use a quite different approach. They think of SP as a characteristic of interaction between a specific event and a specific person. This, of course, means that in order to make a prediction they must independently determine the SP associated with the particular event about which the prediction is made for the person whose behavior is to be predicted. Such a strategy for theorizing has two major difficulties. The first is that it makes little attempt at generality. Obviously some characteristics of an event are important in determining its SP and some are not; we need to know which are which. Furthermore, if SPs are to be used to predict real-world decision making behavior (e.g., for designing displays in military information processing systems), some set of rules which govern the relationship between event and resulting SP must be found so that predictions can be made without running experiments on each situation and person to which the predictions are intended to apply.

What we need and do not have is some systematic empirical information about various displays of probabilities. In the absence of identified stimulus dimensions relevant to SP other than OP, decision theorists are in as awkward a position as psychologists interested in loudness would be in if nothing were known about the physics of sound. Of course this position can only be made worse by ignoring the one stimulus variable which we know to be highly correlated with SP and therefore important: OP. But OP is not enough. Decision theorists cannot much longer get along without empirical information about the psychophysics of probability displays.

Since we have no useful information about displays of probabilities, this paper can only assume that they are important but leave their effects unspecified. The following discussion assumes that each event to which a decision theory might be applied has certain as yet unspecified characteristics which determine its SP. One of those characteristics is, of course, OP, whenever OP is defined. The others are unknown; this paper will call them identifying characteristics. Only two assumptions about identifying characteristics will be consistently made. The first is that whenever an event is repeated as exactly as possible, its identifying characteristics are unchanged; in other words, identifying characteristics are recurrent, not unique. (The contrary possibility would make them inaccessible to scientific study.) The second assumption is that it is quite possible to have events with different identifying characteristics but the same OPs; this simply means that OP is not perfectly correlated with identifying charac-

teristics. (If it were, then Theorem 3 below would imply that SP always equals OP, and so the point of the notion of identifying characteristics would be lost.)

The SP Function Book

Mathematically, this paper will conceive of SP as a function of OP and of identifying characteristics. One way to think about this function is to imagine a book. Each page of this book has on it a page number (the quantity r of the Appendix) and a function which relates SP to OP. These functions need not necessarily be continuous. The page number is simply a code which refers to a particular set of identifying characteristics. Thus, given any event, you can find its SP by finding out its identifying characteristics, looking up the number corresponding to that particular set of identifying characteristics, finding the page in the book which has that page number, and then looking up the SP corresponding to the OP of the event in question on that page. Mathematicians may find this concept and the notation of identifying characteristics clearer if they examine the axiom system and proofs of Theorems 2 and 3 in the Appendix. The assumptions of that proof formally define ASEU models of this class. It turns out that the formulation in terms of a function book, even if that book may have an infinite number of pages, implies that SP always equals OP if additivity of SPs is assumed. But the notion of an SP function book is appropriate for NASEU models, for which no such implication exists.

The idea of the previous paragraph should be familiar to psychophysicists. This model is analogous to a familiar model in auditory psychophysics. Perceived loudness of pure tones is a function of both stimulus intensity and stimulus frequency. A number of different functions relating loudness to intensity, each using a different value of frequency as a parameter, are all an auditory psychophysicist would need to make predictions about loudness. (He would actually use equal loudness contours for the purpose, but the underlying model is the same.)

The purpose of this notion is to provide an orderly way to deal with the possibility that more than one SP may correspond to a given OP, depending on display characteristics. Of course, I hope that such a book of SP functions would have relatively few pages, and an orderly structure, so that the sequence of pages is not arbitrary. But neither of these characteristics need be assumed in the following discussion.

Incidentally, uncertain events as well as risky ones can be provided for within the notion of a book of SP functions. Presumably uncertain events will have identifying characteristics different from risky events (indeed having no definable OP is presumably itself one identifying characteristic), and so pages appropriate for such events can be segregated in one section of the book. Each such page has, as its function relating SP to OP, a flat horizontal line running from $OP = 0$ to $OP = 1$. (More precisely, this line is defined only for $0 < OP < 1$.) Thus once its identifying characteristics are known, the SP of an uncertain event is completely specified.

ASEU MODELS

In spite of major differences in the axiom systems which lead to them and in the philosophy which motivates them, for the purposes of this paper most of the ASEU models are very

similar mathematically and operationally to expected utility maximization models (hereafter called EU models) like the one proposed by von Neumann and Morgenstern (1947). The only difference, and it is a crucial one, is the use of SPs instead of OPs in the models.

Finite vs. Infinite Event Sets

Two main classes of possible ASEU models are defined by the dichotomy between finite and infinite event sets. Mathematically, the distinction is simple enough. If we consider all possible outcomes of 10 tosses of a coin, then this is a finite event set. If we throw a dart at a target, then all possible places it can hit form an infinite event set (note that the point of the dart need *not* be a mathematical point with zero area). Obviously a person can encounter only a finite set of events in his lifetime, but that is irrelevant. The point is that he will encounter many situations in which some one of an infinite set of possible events may occur. At first glance, it would seem desirable, then, to have a decision model capable of coping with any member of an infinite event set—and most decision models are just that. A few theorists (notably Davidson, Suppes, & Siegel, 1957) have preferred to think about finite event sets, mostly because they wanted to be able to check all of their predictions, instead of some finite subset of an infinite set of predictions. As will become clear later, finite event sets immensely complicate the mathematics, and at the same time reduce the value of the model by making it inapplicable to situations in which the set of possible events is infinite. Although this paper discusses finite models below, I consider such models far less interesting than the infinite models.

Please remember that the notion of an infinite event set is not the same thing as the notion of a continuum of OPs. If all elements of a continuum of OPs are realizable by members of an event set, then that set is infinite (nondenumerably infinite, to be precise), but an infinite set of events, all with different OPs, may still have the property that between any two OPs which are associated with members of the event set lies an OP which is not. In fact, continuity of OPs is not assumed directly in Theorems 2 and 3 below, though it can be deduced as a theorem from the nonatomicity property which is assumed.

Restricted vs. Unrestricted Sets of SP Functions

How many different SPs may be associated with different events all having the same OP? Or, to ask essentially the same question in a different form, how many different functions may relate SP to OP? Still another formulation: how many pages may there be in the SP function book? Three classes of answers to this question exist for infinite event sets, and two for finite ones. For infinite event sets, the number of different SPs which may be associated with different events all having the same OP may be finite or infinite. If that number is infinite, it may be denumerable (capable of being set in one-to-one correspondence with the integers) or nondenumerable (the set of all points on a line segment is nondenumerable; so is the set of all OPs). It turns out that the finite and denumerably infinite cases have the same properties, so only the latter, less restrictive case will be considered.

The formulation of the relation between OP and SP in terms of the SP function book tacitly presupposes that the set of different SPs which

may be associated with different events all having the same OP is at most denumerably infinite. If, as the preceding discussion has suggested, the difference between two different events with the same OP but different SPs lies in the way in which that OP is displayed, then the assumption of denumerability is simply the assumption that all possible ways in which an OP could be displayed could be listed (in infinite time). It is difficult to see why anyone would resist so mild an assumption—and yet Theorem 3 will show that its consequences are drastic. Nevertheless, it seems greedy for a theorist to insist that a denumerably infinite set of different SPs to associate with a given OP is not enough and that he wants more. I prefer to accept denumerability and abide by its rather drastic consequences.

For mathematical convenience, however, it is expedient first to examine a restriction wich applies even to the nondenumerable case. It is well known that an infinite number of different probability measures can be defined on a given event set. This makes it seem unlikely that any substantial formal restrictions exist on models of the nondenumerable "anything can happen" sort. But there is at least one. Usually SPs rise with OPs. In fact, it is tempting to assume that they always do so. Consider two events, A and B. If the OP of A is higher than the OP of B and the SP of A is less than the SP of B, call that pair of events an *inversion.* The assumption that SPs always rise with OPs is an assumption that inversions cannot occur. But unless inversions exist in infinite numbers and at all values of OP, then it is easy to prove that SP always equals OP. Theorem 2 below states a stronger assertion from which this one is a self-evident consequence.

Theorem 2: Assume that an infinite, nonatomic event set is being considered, that every event has an SP and an OP, that the SPs of any EEE must add to one, and that some event can be found for which SP $\neq$ OP. It follows that for any OP p, $0 < p < 1$, two events A and B can be found such that $p(A) = p$ and $p(B) < p(A)$ and $\psi(B) > \psi(A) + Q$ where Q is a noninfinitesimal quantity (unless the maximum difference between SP and OP is infinitesimal, or $p(A)$ or $p(B)$ is infinitesimally different from 0 or 1).[2]

A more precise mathematical formulation of Theorem 2, and its proof, are long, technical, and contribute little to the following discussion. I have therefore put them in a mathematical appendix which follows the list of references. The only important assumption necessary to prove Theorem 2 which has not already been discussed is that of nonatomicity; a discussion of it follows Theorem 3, for which it is also essential. Its content is primarily technical. The assumption that all events in the event space have both OPs and SPs has been discussed. The assumption that some event exists for which SP $\neq$ OP is the only reason for even thinking about SPs in risky situations; without it there is no point in a theory about SPs.

Lemma 4 of the mathematical appendix implies some further statements about subjective probabilities in this nondenumerable case. If SP is ever different from OP, there must be some event or events for which the size of that difference (disregarding sign) is a maximum, or at any rate is only infinitesimally different from a maximum. In fact, because of the additivity of SPs there must be at least two such events, one in

[2] This theorem and its proof were formulated by W. M. Kincaid.

which the difference is positive and one in which it is negative. Call the event with the smallest OP for which this maximal difference is present and positive M. Event M has an OP $p(M)$ and an SP $\psi(M)$. For any Event A, define $d(A) = \psi(A) - p(A)$. Then $d(M)$ is the largest positive value of $d(A)$ that can occur, and $d(\bar{M})$ is the largest negative value of $d(A)$ that can occur. Now what can be said about the difference between SP and OP for events other than M? For any such Event A, max $d(A) = d(M)$, and min $d(A) = d(\bar{M})$. Lemma 4 shows that for every probability p there must be an Event B having OP p such that:

$$d(B) = [d(M)]\left[\frac{q(B)}{q(M)}\right]$$

if

$$p(B) \geqq p(M)$$

and

$$d(B) = [d(M)]\left[\frac{p(B)}{p(M)}\right]$$

if

$$p(B) \leqq p(M)$$

There must also be an Event C having OP p such that:

$$d(C) = [d(\bar{M})]\left[\frac{q(C)}{p(M)}\right]$$

if

$$p(C) \geqq q(M)$$

and

$$d(C) = [d(\bar{M})]\left[\frac{p(C)}{q(M)}\right]$$

if

$$p(C) \leqq q(M)$$

Finally, it shows that events must exist having that same OP p and all possible values of ψ between $\psi(B)$ and $\psi(C)$. In other words, in a space relating SP to OP there is a parallelogram of points defined by the four vertices (0, 0) (1, 1), $[p(M), \psi(M)]$, and $[p(\bar{M}), \psi(\bar{M})]$. All points within this parallelogram must be realizable by events in the event space with appropriate subjective probabilities.

The preceding paragraph leads up to Theorem 3. The awkward property exhibited in Theorem 2 applies to models which permit a nondenumerably infinite set of SPs to be associated with any given OP. Next consider what happens if any OP can be found (other than 0 or 1) with which no more than a denumerably infinite set of SPs is associated.

Theorem 3: As in Theorem 2, assume an infinite, nonatomic event set, each event having an SP and an OP, and assume that the SPs of any EEE must add to one. In addition, assume that there is some OP other than 0 or 1 which has the characteristic that the number of different SPs which occur in conjunction with different events all having that OP is at most denumerably infinite. It follows that SP always equals OP.[3]

Again a more precise formulation of the theorem, and its proof, are in the mathematical Appendix. Certain additional assumptions of purely technical importance (e.g., continuity at zero) are necessary to the proof, but all the assumptions which I consider at all controversial are stated here.

At this point it is worth while considering the assumption of nonatomicity which underlies Theorems 2 and 3. Suppose I assert "For any event which is already included in an event set, I wish to include another which is defined as the simultaneous occurrence of the event in question and heads in the toss of a coin which I happen to have handy." This assertion defines a nonatomic event set (in fact, it is somewhat stronger than

[3] This theorem and its proof were formulated by L. J. Savage. W. M. Kincaid turned Savage's sketched proof into the form contained in the Appendix.

is necessary to specify a nonatomic event set). In other words, the property of nonatomicity simply asserts that any given subset may be further subdivided. Unless the set of events with which you intend to work is rigidly defined in advance, it is impossible, as I see it, to think of any real world event set which is not nonatomic. However, this property is far from trivial. For example, the assumption of nonatomicity directly implies that every OP is realizable by some event in the event set, and that the event set is nondenumerably infinite.

All in all, the characteristics of ASEU models based on infinite classes of events seem unattractive. We may adopt any of three strategies. One is to accept and live with these unfortunate properties. That possibility has already been examined. It is now time to mention very briefly the second set of possibilities.

Finite Event Sets

The arguments given previously against finite events sets were, I think, strong enough to argue against their serious consideration. However, it is interesting to speculate further about their mathematical properties. It seems quite likely to me that analogues of Theorems 2 and 3 must exist for finite cases. Such theorems would define maximum differences between SP and OP as a function of number of events in the finite event set, spacing of the OPs of those events on the OP continuum, number of different SP functions permitted (if that number is less than the number of events in the event set), and number and location of inversions of ordering between SPs and OPs. It seems plausible both that such theorems exist and that Theorems 2 and 3 are limiting cases as the number of events in the event set approaches infinity. Unfortunately, I have been unable even to guess what the form of such theorems might be (except for specially constructed cases, such as the case of equal spacing of events along the OP continuum). As is often the case, it looks as if the finite cases are mathematically much more difficult than the infinite cases. Even in the case where only a single monotonic function relating SP to OP is permitted (surely the simplest finite case), I have been unable to formulate a specific theorem relating number and spacing of events in the event set to maximum possible difference between OP and SP.

The Third Strategy

This paper has considered two classes of ASEU models based on infinite event sets and a class based on finite event sets; all are unattractive for various reasons. The third possibility is to abandon ASEU models altogether and use NASEU models instead. Some previously discussed considerations about utility functions make this option attractive. But the NASEU models are so much less powerful than the ASEU models that I would be very reluctant to turn to them on a priori grounds alone. So before examining them seriously, it is worth while to review the empirical evidence which bears on them.

THE PROBABILITY PREFERENCE DATA

A rather substantial body of data bears on the relation between SP and OP. These data come from a series of experiments (Edwards, 1953, 1954a, 1954b, 1954d, 1955) which I have collectively called the probability preference experiments.

The basic design of the probability

preference experiments was as follows: Lists of bets were prepared. Each list included eight bets, with OPs of winning (or losing, depending on the list) running from 1/8 through 8/8. In various experiments a total of eight different lists were used. The eight bets in each list all had the same objective EV, so that there was no objective reason to prefer one bet in a given list to any of the others in the same list. All bets in each list were paired with one another according to the method of paired comparisons. At various times, a total of well over 100 college students were required to choose between the members of these pairs. In some experiments, the *S*s sat in classrooms, looked at slides of the pairs of bets, and made choices "as if they were gambling." Other *S*s, run individually, first made imaginary choices, then gambled for worthless chips, and finally for real money.

The main results of these experiments, taken as a whole, were:

1. Although there were substantial individual differences in choices, certain patterns of choices showed up in all experimental groups and in just about all individuals. The two most outstanding of these patterns were that *S*s usually preferred the bet with the 4/8 probability of winning from any positive EV list to the other bets on that list, and usually preferred the bet with the lower probability and higher amount of loss in any pair of negative EV bets.

2. The preferences cited above, and indeed the complete pattern of choices, were relatively independent of EV level of the list of bets involved, so long as the zero point was not crossed. This means that the preferences observed should be attributed to the OPs involved, which were constant, rather than to the amounts of money, which varied from list to list. This fact suggests that for these students and these amounts of money, the utility curve is relatively linear with amounts of money, while the SP curve is not linear with OP. It also suggests that the value of OP from which SP deviates most widely is 0.5, for positive EV bets. Using the assumption that SP = OP and the further assumption that the size of the just noticeable difference for utility is half a cent, it is impossible to construct a utility function to account for the preferences observed. If the size of the just noticeable difference for utility is assumed to be zero, of course such a curve can be constructed, but it has at least 12 inflection points between \$0 and +\$6. The same sorts of statements can be made if the data are analyzed *S* by *S*.

3. The complete pattern of choices changed radically from positive EV to negative EV bets, even though exactly the same OPs were used. The main change was a strong preference for negative EV bets in which the probability of losing was low and the amount of possible loss high. When this preference was removed from the data by crude statistical means, the residual preference pattern was pretty much the mirror image of the preference pattern for positive EV bets. These findings suggest that there is a strong interaction between utility and SP.

4. The differences among just imagining, gambling for worthless chips, and gambling for real money are discernible, but not sufficiently serious to invalidate any of the conclusions discussed above. This means that the difficulties which might arise in a gambling experiment because *S* wins or loses on each bet before he decides about the next bet are much less severe than a priori considerations might lead one to expect.

5. Minor changes in wording, in the choice of payoff events and in the financial status of *S*s do not make any detectable differences in their choices. In particular, *S*s can be made to win exceedingly large amounts of money, or made to lose substantial but smaller amounts of money, without significantly changing their patterns of choices. This finding has been confirmed in a study devoted especially to it (Edwards, 1954d).

Findings 1 and 2 cause me to reject the hypothesis that SP = OP. Therefore, they, in conjunction with the considerations discussed in the previous section, cause me to reject all EU and ASEU models. The implications of Findings 3, 4, and 5 for a NASEU model are important; they will be discussed below.

NASEU Models

What are the mathematical properties, if any, of a NASEU model? Take a look at Equation 1. It defines NASEU as well as ASEU models. But what does the operation of addition called for by the sigma in Equation 1 mean for a NASEU model? For an ASEU model, that sigma simply derives from the usual definition of a probability measure, and reflects the additivity property of that measure. For a NASEU model, it has no such simple mathematical justification. A variety of psychological arguments about its plausibility are available. This paper will not press them; evidence to be presented later argues against this form of NASEU model, and I do not advocate or defend such a model. The purpose of this section is to introduce some ideas which are a part of the model which I do defend. For the time being, therefore, it is enough to say that no law of men or mathematics forbids the multiplication of utility and nonadditive SP values, the algebraic addition of the products, or test of the hypothesis that people make decisions in such a way that they in fact maximize the resulting sums.

Ratio Scale Utilities

However, Theorem 1 indicates that a NASEU model must use a ratio scale measure of utility. A ratio scale of utility implies that utility has a true zero point. Where is it? Only one answer is plausible: where you now are. Zero utility is your current position, and you can never leave it.

This is not exactly a new idea. The first implicit use of it was by Mosteller and Nogee (1951), who used the EU model to make a utility scale for money (and, incidentally, proposed nonadditive SPs as an alternative explanation for their results). Mosteller and Nogee used a gambling situation in which real money changed hands during the course of the experiment. But in determining the utility of a given amount of money, they simply found an indifference point which involved that amount of money without moving the origin to take into account *S*'s financial status at the time they measured it. Such a procedure can yield a classical utility function only if the form of that function is invariant up to a linear transformation under movements of the origin (through which all utility functions conventionally pass) along the function. Only a very limited class of functions, of which the most familiar member is the straight line, has this property, and the utility curves Mosteller and Nogee obtained don't look like any member of that class. Consequently those curves in fact defined zero utility as the present monetary status of each *S*, and so

were of the type that this section advocates.

The first self-conscious use of an idea like that proposed here was made by Markowitz (1952). He suggested that the zero point of the utility scale be taken as *S*'s customary financial position, and that the form of the function changes when that customary financial position changes. He used this idea to remedy a deficiency in Friedman and Savage's (1948, 1952) previous account of gambling and insurance buying. The only difference between Markowitz's position and this one is that Markowitz defines zero utility as the customary position, while this paper defines it as the current position.

Although this conception of utility is novel and quite different from the traditional one, it is neither internally contradictory nor absurd. Nor does it imply that people cannot change their habits as they change their financial status. Such changes in habits simply show up as changes in the shape of the utility curve. Furthermore, this point of view is not empirically contentless. If it is to be useful, people must not make serious changes in their subjective value scale for a commodity as a result of relatively small changes in the amount of that commodity which they possess. This proposition is in principle testable, though testing it experimentally is difficult. However, Findings 4 and 5 of the probability preference experiments (see above) suggest that it may be correct.

Indifference Curves and a Property of NASEU Models

Loss of the addition theorem for SPs, even though compensated by a true zero for utility, is a very serious loss. Can a model of the type defined by Equation 1 mean much without an addition theorem for SPs? Yes. Equation 1 implies a strong decomposability property which has considerable significance both mathematically and empirically. The following discussion exhibits this property and derives two theorems, one of which can be empirically tested. These theorems have the additional advantage of being true for all plausible models defined by equations like Equation 1. This means that an empirical verification of them would be strong evidence for the usefulness of the class of maximization models with which this paper is concerned, while an empirical disproof would necessitate search for a quite different kind of theory of decision making (e.g., one based on variance preferences).

The following discussion deals exclusively with bets of the following form: Event E has probability p of occurring. If it occurs, S receives an amount of money x, which is greater than 0. If it does not occur, no money changes hands. Any such bet can be described as a point (x, p); that is, all such bets can be completely represented by a plane bounded at $p = 0$, $p = 1$, and $x = 0$. Now, consider the bets (x, p) and $[(x + e)p]$ where e is a small positive number. A great deal of evidence, plus common sense, suggests that most *S*s would prefer the second of these two bets to the first. Similarly, $[x, (p + f)]$ is preferable to (x, p) where f is another small positive number. If both of these assertions continue to be true as e and f approach zero for any values other than zero of x and p (a reasonable idealization), and if the direction of preference reverses for any negative values of e and f (another reasonable idealization), it follows that a set of indifference curves can be drawn in the xp plane. An indifference curve in this application is simply a function which defines a

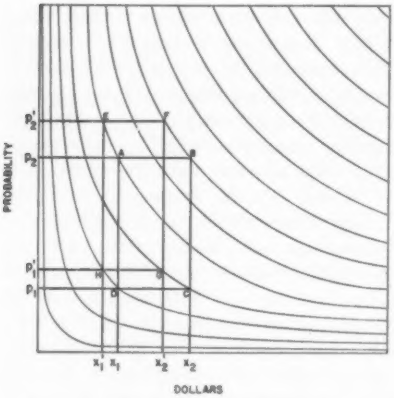


FIG. 1. Hypothetical indifference map among simple bets which have "no money changes hands" as one of the two possible outcomes. The rectangles are described in Theorems 4 and 5 of the text.

set of bets among all of which S is indifferent (see Coombs & Beardslee, 1954; Edwards, 1954c for details). A set of indifference curves is called an indifference map; Figure 1 shows such a map. The properties to be discussed below are general properties of a certain family of indifference curves and have applications to the use of indifference curves in economic theory; these applications will be discussed at the end of this section.

What does the SEU model have to say about such indifference curves? Since each curve identifies bets among all of which S is indifferent, all bets on a given curve must have the same SEU. In short, in utility and SP measures the equation of each indifference curve c must be:

$$\psi_c u_c = k_c \qquad [4]$$

Of course there is a different value of k for each curve.

But indifference curves are stated in physical, not subjective, units. What implications does Equation 4, and therefore SEU maximization models, have about the form of indifference curves stated in physical units? It is obvious that if any single indifference curve is selected, it is possible to find a transformation on the probability and money axes which will make it conform to Equation 4, which is of course the equation of a rectangular hyperbola. In fact, any set of two indifference curves could be made to conform by proper choice of transformations. But that uses up all the arbitrariness implied in Equation 1. To put it another way, the transformations necessary to make these two curves conform to Equation 4 completely specify the utility and subjective probability functions for this S. Now, if Equation 4 is to be correct, it must turn out that all other indifference curves in the xp plane must be turned into rectangular hyperbolas by the same utility and SP functions which turned the first two into rectangular hyperbolas. This is an exceedingly strong condition on the family of indifference curves. The following proof shows that an indifference map satisfies this requirement if and only if a quite specific relationship exists among the slopes of the indifference functions passing through any four points at the corners of a rectangle whose sides are parallel to the coordinate axes.

The theorem assumes that an indifference map is given each of whose functions is continuous, monotonic decreasing, and has continuous derivatives. The utility and SP transformations are assumed strictly monotonic increasing. They are also assumed single-valued, which means that the events displaying the OPs used must all have the same identifying characteristics if Theorems 4 and 5 are to be applicable. Two values of money, x_1 and x_2, and two values of probability, p_1 and p_2, are chosen at random (or systematically; it

doesn't matter how they are chosen so long as both amounts of money are greater than zero and both probabilities lie between zero and one); they of course jointly define four bets. Let us call these four bets A, B, C, and D, as in Figure 1. From now on, S_A means the slope of the indifference curve which passes through Bet A.

Theorem 4: If any kind of SEU model, additive or otherwise, is correct, then $S_A S_C = S_B S_D$, and if this condition on the slopes of an indifference map is satisfied, and if at least one indifference function has a negative slope, then it follows that for these data a NASEU model or some more powerful SEU model is correct.[4]

The proof of Theorem 4 is in the mathematical appendix.

The property of all SEU maximization models, including NASEU models, which is established in Theorem 4 is theoretically elegant and in principle empirically testable. No such empirical test is likely to take place, however; differentiation of empirical curves requires more precision than is usually obtainable in experiments using human *S*s. Fortunately, the microscopic property described in Theorem 4 has a macroscopic equivalent, which can be tested experimentally.

Look at Figure 1 again. Choose an amount of money x'_1, significantly smaller than x_1. Extend a vertical line from x'_1 to the indifference curve which passes through D; call the intersection point H. From H extend a vertical line until it intersects the indifference curve which passes through A, and a horizontal line until it intersects the indifference curve which passes through C; call the two points thus identified E and G. A horizontal line can be extended from E and a vertical one from G, intersecting in a point F and thus completing another rectangle. Will Point F lie on the indifference curve which passes through Point B? It turns out that it will if and only if the indifference map has the property expressed in Theorem 4. It is probably intuitively apparent that the property of Theorem 4 can be derived from this one; Theorem 4 is simply the limiting case as the difference between x_1 and x'_1 becomes infinitesimal. This property can be expressed formally as

Theorem 5: If any kind of SEU model, additive or otherwise, is correct, then if a second rectangle is constructed as in Figure 1 and three of its vertices are chosen to lie on the same indifference curves as three of the vertices of the original rectangle, then the fourth vertex of the new rectangle must lie on the same indifference curve as the fourth vertex of the original rectangle. Furthermore, if the vertices of the new rectangle are connected by straight lines to the vertices of the old rectangle, the slopes of the connecting lines will be related by the equation $S_{\overline{AE}} S_{\overline{CG}} = S_{\overline{BF}} S_{\overline{DH}}$, where $S_{\overline{AE}}$ symbolizes the slope of the line connecting Point A with Point E in Figure 1. If these conditions are satisfied, and if at least one indifference function has a negative slope, then it follows that for these data a NASEU model or some more powerful SEU model is correct.[5]

To prove Theorem 5, start by looking again at Equation 4. With notation taken from Figure 1, it says that:

$$g(x'_1)h(p'_1) = g(x_1)h(p_1) \quad [5]$$

$$g(x'_1)h(p'_2) = g(x_1)h(p_2) \quad [6]$$

$$g(x'_2)h(p'_1) = g(x_2)h(p_1) \quad [7]$$

[4] This theorem and its proof were formuated by George Minty.

[5] This extension of Minty's theorem was formulated by L. J. Savage.

where $g(x_1)$ is the utility of x_1, $h(p_1)$ is the SP of p_1, and so on. Now multiply the left half of Equation 6 by the left half of Equation 7 and divide by the left half of Equation 5, and similarly for the right halves of these equations. The results are:

$$g(x'_2)h(p'_2) = g(x_2)h(p_2) \qquad [8]$$

Equation 8 is what the first half of Theorem 5 asserts.

Now note that if a line is drawn from A to E and similarly for the other three pairs of vertices, each of these lines is a chord of an indifference curve. The slope of any of these lines is given by:

$$S_{ij} = \frac{p'_j - p_j}{x'_i - x_i} \qquad [9]$$

where:

$$i, j = 1, 2$$

A little algebra will show that $S_{AE}S_{CG} = S_{BF}S_{DH}$, which is what the second half of Theorem 5 asserts. This result is a close macroscopic analogue of Theorem 4. In fact, Theorem 4 follows immediately from it if x'_1 is allowed to approach x_1, for then the four chord slopes approach the slopes of the indifference curves at the vertices of the original rectangle. This proof of Theorem 5 is only a proof of necessity; a proof of sufficiency very similar to that given in the Appendix for Theorem 4 exists, but will not be given here.

The property of NASEU (and other) models expressed in Theorem 5 can be tested by trying to perform the construction specified in the first half of Theorem 5. The relationship among slopes specified in the second half of Theorem 5 is tautologously true for all pairs of parallelograms with sides parallel to the coordinate axes, and so provides no additional information concerning the NASEU model. An experiment designed to test this property, and so to test this class of models, is now under way.

Although the foregoing discussion was expressed in terms of bets, it is actually a quite general property of certain kinds of indifference curves. The properties of Theorems 4 and 5 clearly apply to any indifference map whose curves have the form $f(x)g(y) = k$. But they also apply to any indifference map whose curves have the form $f(x) + g(y) = k$. (To prove this, simply replace the multiplications in Equations 5–7 by additions and the divisions by subtractions. The proof now proceeds as before, using additions instead of multiplications and subtractions instead of divisions up to but not including Equation 9. A similar modification of the proof of Theorem 4 is possible.) This additive form of the indifference curves corresponds to an indifference map concerned with two goods which are independent, rather than competing (like butter and margarine) or completing (like right and left shoes). Thus Theorems 4 and 5 provide an operational test of the hypothesis that two goods are independent of each other for a given indifference map—so far as I know, the only operational test so far suggested which can be applied directly to an indifference map. Of course this test is applicable only if the existence of at least interval scale utilities is assumed.

The Weighted SEU Model

Most of the argument of this paper has been devoted to the proposition that for any of the class of models considered here except NASEU models any function relating SP to OP may only be linear. The probability preference data indicate that it cannot be linear. Or do they?

An old familiar finding in psycho-

physics is that the form of any subjective scale depends on the methods used to determine it. The same proposition may be true for SP and utility scaling. The probability preference experiments were a rather indirect method inferring utility and SP properties from gambling choices. What results would more direct methods yield?

In the case of utility of money, it seems unlikely that direct methods of psychophysical scaling will be useful. The kind of question which would have to be asked would sound like "What amount of money is twice as valuable as two dollars to you?" It seems unlikely that any *S* would give an answer systematically different from "Four dollars," no matter what his utility function for money might look like. The difficulty is that money has an invincibly numerical character, and most people would probably respond to mathematical properties of these numbers rather than to any subjective values they might have. In any case, the evidence is continuing to accumulate that for small amounts of money utility is linearly related to dollar value—in which case psychophysical determination of utility scales is not very likely to be interesting. Of course psychophysical methods could be used to scale the utility of other objects which have a less numerical character, but it isn't obvious how such experiments would shed much light on general properties of utility functions, because in such experiments the concept of objective value is difficult to define.

In the case of SP, on the other hand, some psychophysical data exist. Shuford (1959) performed an experiment in which he showed *S*s square matrices made of 400 thumbtacks, some red and some green. The *S*s were required to estimate the relative frequency of each kind of thumbtack in the display after a brief look, and succeeded in estimating with remarkably good accuracy. Relative frequency in a matrix isn't exactly the same thing as OP, but is very closely related. Shuford's experiment suggested one in which I used these square matrices as the basis for probabilities in the probability preference experiment situation. I displayed two such matrices and associated a bet with each. The *S*s were required to estimate the relative frequency of each color in each of the two matrices, and then were required to choose one of the two bets. A random device (two 20-sided dice) chose one location on the matrix associated with the chosen bet, and the color of the thumbtack at that location determined the outcome of the bet. Preliminary results show that the estimates of relative frequency are excellent, just as in Shuford's experiments, but that the choices made among the bets are distorted. In short, at almost the same moment the *S*s say that SP is and is not equal to OP. Something peculiar is clearly going on.

My interpretation of this finding is that people perceive OPs correctly, but that they misuse them. They simply prefer to bet at some OPs rather than others, even though they perceive these OPs quite accurately. This hypothesis can be translated into mathematics by supposing that they attach to each event a weight w_i which expresses the relative desirability or undesirability of the probability displayed by that event. A model of this sort is only a minor variation on Equation 1:

$$\text{WSEU} = \sum_i (p_i w_i) u_i \qquad [10]$$

Mathematically, the WSEU model is exactly the same as the NASEU

model (which is why this paper has devoted so much attention to the properties of the NASEU model) except that the WSEU model comes by its addition operation naturally rather than arbitrarily. Since there is no mathematical difference between the NASEU and WSEU models, the difference between them is presumably one of esthetics; I find the WSEU model esthetically more appealing.

I cannot resist speculating a bit about the nature of a WSEU model. I assume, of course, that it applies to infinite event sets. How many different sets of weights are necessary? In the language used earlier in this paper, how many different pages are there in the SP book? The data now available suggest the speculation that there may be exactly five pages in that book, each page defined by a class of possible payoff arrangements. In Class 1, all possible outcomes have utilities greater than zero. In Class 2, the worst possible outcome (or outcomes, if there are several possible outcomes all with equal utility) has a utility of zero. In Class 3, at least one possible outcome has a positive utility and at least one possible outcome has a negative utility. In Class 4, the best possible outcome or outcomes has a utility of zero. And in Class 5, all possible outcomes have negative utilities. (Note that this classification is meaningful only because of ratio scale utilities.) This speculation is partly based on Finding 3 of the probability preference experiments already discussed.

Another speculation of some interest concerns possible limitations on the complete freedom in choice of weights in WSEU models (or in choice of SPs in NASEU models). It is extremely unlikely, for example, that anyone would have a set of weights or a SP function which assigned Wp (or ψ) = 0.9 both to the occurrence and to the nonoccurrence of an event A. Although people do not use additive SP functions, at the same time they do not behave in as lawless a fashion as the discussion has so far suggested. The problem is to know what the actual limitations on additivity are, and this problem is primarily empirical. Information can be gained by inspection of SP functions, by examination of the results of experiments in which *S*s must combine probabilities, and perhaps in other ways also. Several axioms which capture parts of this idea of approximate additivity can be dreamed up; it is a question whether anything useful can be deduced from them (so far nothing has been):

1. $\psi(A) > 0.5$ if and only if $\psi(\bar{A}) < 0.5$.
2. If A_i is the i^{th} element of a partition of S, then $\left|\sum_i A_i - 1\right| \leq k$ where k is some fixed constant less than 1.
3. If A_i is the i^{th} element of an n-fold partition of S, then $\left|\sum_i A_i - 1\right| \leq f(n)$, where $f(n)$ is nondecreasing and has a maximum value or a limit.

Note that these possible axioms are alternative formulations, not parts of the same structure. Whether they or anything like them are of any use depends on what (if anything) they imply, and on the data. Axiom 3, which states that the extent of the failure of additivity depends on the number of elements in the partition of S, is particularly attractive to me because of some unpublished data which suggest its usefulness.

Of course, all the models considered in this paper are static. That means that they do not consider possible changes over time. Static models are presumably simpler and easier to study than dynamic ones, which do

consider changes in time. The usual real life decision situation, however, requires a man to make not one but a sequence of decisions. These decisions are embedded in a flow of incoming formation. They change the nature of the incoming information in two ways, by modifying the environment which generates it and by modifying the relation between environment and person which makes it available. Thus the crucial question for most real-life decisions concerns the relation between information flow and decision making. But that is the problem of dynamic decision making, which is not the subject of this paper.

Variance Preferences

Allais (1953a, 1953b) and others both before and after him have suggested that the variance of a bet may be more important than its EV, objective or subjective, in determining preferences. This suggestion is intuitively appealing, but extremely hard to translate into an experiment. I published one attempt to establish the existence of variance preferences (Edwards, 1954d) in which the conclusion was that they exist, but are small in size relative to probability preferences. Coombs and Pruitt (1960) have some data which suggest that both variance and skewness preferences exist, and are important. In any case, the existence of variance or skewness preferences is likely to be inconsistent with any model like Equations 1 or 10, or indeed with any maximization model of the kind discussed in this paper. Unfortunately, the Coombs and Pruitt data were choices among two-outcome bets. For such bets, variance, skewness, payoffs, probabilities, and expected value all are related by a network of equations which mean that there are too few degrees of freedom to ascertain with which among these quantities preferences are in fact correlated. The only way I know to disentangle all these quantities empirically is to use bets with more than two possible outcomes; both Coombs and I plan such experiments.

The conditions on slopes of indifference curves which were stated in Theorems 4 and 5 would in principle provide experimental tests for the reality of variance and skewness preferences, in the sense that if either is satisfied then variance and skewness preferences cannot have played a role in the choices from which the indifference curves were constructed except as they were reflected in utility and SP functions.

If, as seems likely, experiments reveal that variance preferences exist, and are not reducible to utility and SP functions, what direction should theoretical endeavor take then? I don't know. One thing is clear: such a demonstration would be the final blow for the class of models with which most theorists of decision making have been concerned until now and with which this paper has dealt.

SUMMARY

Most contemporary decision theories explain choices among risky alternative courses of action by supposing that the available course of action with the highest subjectively expected utility will be preferred to all others. Subjectively expected utility is calculated by finding the subjective value or utility of each possible outcome of the course of action, multiplying it by its subjective probability of occurrence, and summing these products over all possible outcomes. The concept of subjective probability in this theory is quite confusing.

This paper shows that if the subjective probabilities of a set of events of which some one must happen and not more than one can happen must add up to a specified constant (usually taken as one for convenience), then utility can be measured on an interval scale and subjective probability on a ratio scale. But if subjective probabilities do not have this additivity property, then both utility and subjective probability must be measured on a ratio scale.

If subjective probabilities do have the additivity property, if the decision model in which they are used is in principle applicable to any conceivable set of events, and if the number of different subjective probabilities which may occur in conjunction with a given objective probability is no more than denumerably infinite, then whenever objective probabilities are defined subjective probability must equal objective probability. This theorem and another mean that the concept of subjective probabilities which must add to one, and decision models based on that concept, cannot be very helpful in explaining decision making; a large amount of data indicates that subjective probability cannot always equal objective probability.

One alternative is to use a decision model in which subjective probabilities need not add up to any particular constant. Such a model is possible; its properties are explored. One property is that in such a model, utility must have a true zero point. The only reasonable zero point for utility is where you now are; this notion is not new to decision theory.

Any form of the expected utility maximization hypothesis (with additive or nonadditive subjective probabilities) implies a strong relationship among the slopes of indifference curves for simple bets. This relationship leads to an experimental test of the applicability of models of this class. If the result of this test should be negative, further theories about decision making in risky situations should probably include the concept of variance preferences.

REFERENCES

ALLAIS, M. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école americaine. *Econometrica*, 1953, **21**, 503–546. (a)

ALLAIS, M. La psychologie de l'homme rationnel devant le risque: La théorie et l'expérience. *J. Soc. Statist. Paris*, 1953, **94**, 47–73. (b)

COOMBS, C. H., & Beardslee, D. On decision-making under uncertainty. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes*. New York: Wiley, 1954.

COOMBS, C. H., & PRUITT, D. G. *A study of decision making under risk. Willow Run Lab. Rep.* 1960, No. 2900-33-T. (Univer. Michigan.)

DAVIDSON, D., SUPPES, P., & SIEGEL, S. *Decision-making: An experimental approach.* Stanford. Stanford Univer. Press, 1957.

DVORETSKY, A., WALD, A., & WOLFOWITZ, J. Elimination of randomization in certain statistical decision procedures and zero-sum two-person games. *Ann. math. Statist.*, 1951, **22**, 1–21. (a)

DVORETSKY, A., WALD, A., & WOLFOWITZ, J. Relations among certain ranges of vector measures. *Pacific J. Math.*, 1951, **1**, 59–74. (b)

EDWARDS, W. Probability-preferences in gambling. *Amer. J. Psychol.*, 1953, **66**, 349–364.

EDWARDS, W. Probability preferences among bets with differing expected values. *Amer. J. Psychol.*, 1954, **67**, 56–67. (a)

EDWARDS, W. The reliability of probability preferences. *Amer. J. Psychol.*, 1954, **67**, 68–95. (b)

EDWARDS, W. The theory of decision making. *Psychol. Bull.*, 1954, **51**, 380–417. (c)

EDWARDS, W. Variance preferences in gambling. *Amer. J. Psychol.*, 1954, **67**, 441–452. (d)

EDWARDS, W. The prediction of decisions among bets. *J. exp. Psychol.*, 1955, **50**, 201–214.

EDWARDS, W. *Subjective probability in decision theories.* *Willow Run Lab. Rep.*, 1959, No. 2144-361-T. (Univer. Michigan.)

EDWARDS, W. Measurement of utility and subjective probability. In H. Gulliksen & S. Messick (Eds.), *Psychological scaling: Theory and applications.* New York: Wiley, 1960.

EDWARDS, W. Behavioral decision theory. *Annu. Rev. Psychol.*, 1961, **12**, 473–498.

FRIEDMAN, M., & SAVAGE, L. J. The utility analysis of choices involving risk. *J. pol. Econ.*, 1948, **56**, 279–304. (Reprinted with minor changes in G. J. Stigler & K. E. Boulding [Eds.], *Readings in price theory.* Chicago: Richard D. Irwin, 1952.)

FRIEDMAN, M., & SAVAGE, L. J. The expected-utility hypothesis and the measurability of utility. *J. pol. Econ.*, 1952, **60**, 463–475.

HALMOS, P. R. *Measure theory.* New York: Van Nostrand, 1950.

LUCE, R. D. On the possible psychophysical laws. *Psychol. Rev.*, 1959, **66**, 81–95.

LYAPOUNOV, A. Sur les fonctions-vecteurs complètement additives. *Izv. Akad. Nauk SSSR Ser. Math.*, 1940, **4**, 465–478.

MARKOWITZ, H. The utility of wealth. *J. pol. Econ.*, 1952, **60**, 151–158.

MOSTELLER, F., & NOGEE, P. An experimental measurement of utility. *J. pol. Econ.*, 1951, **59**, 371–404.

SAVAGE, L. J. *The foundations of statistics.* New York: Wiley, 1954.

SHUFORD, E. H. *A comparison of subjective probabilities for elementary and compound events.* Chapel Hill, N. C.: University of North Carolina, Psychometric Laboratory, 1959. (Report No. 20)

VON NEUMANN, J., & MORGENSTERN, O *Theory of games and economic behavior.* (2nd ed.) Princeton: Princeton Univer. Press, 1947.

(Received August 23, 1960)

APPENDIX[A1]

This appendix contains precise statements of the assumptions and proofs of Theorems 2, 3, and 4 from the body of the article. The proofs of Theorems 1 and 5 are contained in the text.

Theorems 2 and 3 assume the existence of a probability space S with events $A, B, C, \cdots$, having objective probabilities $p(A), p(B), p(C), \cdots$; the measure p has the conventional properties. Associated with each Event A is an element $r(A)$ of a space R; $r(A)$ is called the identifying characteristic of Event A. Two or more events (indeed, large classes of events in the models discussed in this paper) may have the same identifying characteristics. A function $\Gamma(r, p)$ is defined for each r and p such that for some Event A, $r(A) = r$ and $p(A) = p$. The subjective probability $\psi(A)$ of the Event A is then defined by the equation:

$$\psi(A) = \Gamma[r(A), p(A)]$$

Several assumptions, mostly of technical importance only, are necessary. They are stated here with little discussion; the important ones are discussed in the text.

Axiom 1: p is completely additive.

Axiom 2: S is nonatomic with respect to p. This means that if $p(A) > 0$, then there exists some Event B contained in A (a subset of A) such that $0 < p(B) < p(A)$.

Axiom 3: For each r for which $\Gamma(r, 0)$ is defined, $\Gamma(r, 0) = 0$, and $\lim_{p \to 0} \Gamma(r, p) = 0$ uniformly in r. That is, given $\epsilon > 0$ there exists a $\delta > 0$ such that $\Gamma(r, p) < \epsilon$ whenever $p < \delta$ and $\Gamma(r, p)$ is defined.

Axiom 4: $\psi(A)$ is a probability measure. Specifically: (*a*) For $A \subset S$, $0 \leq \psi(A) \leq 1$, (*b*) $\psi(S) = 1$ and $\psi(\bar{S}) = 0$, and (*c*) If $A, B, \subset S$ and $A \cap B = 0$, then $\psi(A \cup B) = \psi(A) + \psi(B)$.

There is some redundancy in the statement of Axiom 4; it was included for clarity and to permit Theorem 4 to be proved without the use of Part *c* of the axiom.

Some lemmas prepare the way for the proof of Theorems 2 and 3.

Lemma 1: The probability measure ψ is completely continuous with respect to p, and completely additive.

From Axiom 3, it follows immediately that to any number $\epsilon > 0$ there corresponds a number $\delta > 0$ such that $\psi(A) < \epsilon$ for any Event A such that $p(A) < \delta$. This is what complete continuity means.

[A1] See body of article for credit footnotes identifying the formulators of each theorem.

Now let $A_1, A_2, A_3, \cdots$ be a countable sequence of disjoint events such that $\bigcup_{n=1}^{\infty} A_n = A$. Since p is completely additive, we have $p(A) = \sum_{n=1}^{\infty} p(A_n)$ and (setting $B_n = \bigcup_{k=n}^{\infty} A_k$) $\lim_{n\to\infty} p(B_n) = 0$. It follows from the foregoing discussion that $\lim_{n\to\infty} \psi(B_n) = 0$. Since $p(A)$ exists, so does $\psi(A)$, and since it follows from Axiom 4 that $\psi(A) = \sum_{k=1}^{n-1} \psi(A_k) + \psi(B_n)$ for all n, we conclude that $\psi(A) = \sum_{n=1}^{\infty} \psi(A_n)$, and thus that ψ is completely additive

Lemma 2: S is nonatomic with respect to ψ.

Proof: Lemma 1 makes it possible to apply the Radon-Nikodym theorem (see for example Halmos, 1950, p. 128) to infer that ψ has a probability density with respect to p. If A is any event with $\psi(A) > 0$, it must include a Subevent A' such that $p(A') > 0$ and the density of ψ with respect to p is positive on A'. Since S is nonatomic with respect to p by Axiom 2, there is an event $B \subset A'$ such that $0 < p(B) < p(A')$; consequently also $0 < (B) < \psi(A') \leqq \psi(A)$. This proves the lemma.

Lemma 3: Let f be a (Borel-measurable) function defined on S such that $0 \leqq f(w) \leqq 1$ for w in S, and let Ω_1 and Ω_2 be completely additive nonatomic probability measures. Then there exists an Event E such that:

$$\int_S f(w)d\Omega_1(w) = \Omega_1(E),$$

$$\int_S f(w)d\Omega_2(w) = \Omega_2(E)$$

This result is a specialization of one established by Dvoretsky, Wald, and Wolfowitz (1951a, 1951b), representing an extension of an earlier theorem of Lyapounov (1940).

It is convenient at this point to interpret the objective and subjective probabilities of an Event A as the rectangular coordinates of a point $[p(A), \psi(A)]$ in a Euclidean plane. Then if p and ψ are not identical measures there must exist an Event H for which $\psi(H) \neq p(H)$, and the points $(0, 0)$, $[p(H), \psi(H)]$, $[p(\bar{H}), \psi(\bar{H})]$, and $(1, 1)$ (where $\bar{H}$ is the complementary event to H) form a parallelogram. The following lemma asserts that there is at least one event corresponding to every point within the parallelogram.

Lemma 4: Let H be any event, and let a and b be any two numbers such that $0 \leqq a \leqq 1$ and $0 \leqq b \leqq 1$. Then there exists an Event C such that $p(C) = ap(H) + b[1 - p(H)]$ and $\psi(C) = a\psi(H) + b[1 - \psi(H)]$.

Proof: Let g be the characteristic function of H (that is, let $g(w)$ be 1 for w in H, and 0 otherwise). Then by Axioms 1 and 2 and Lemmas 1 and 2 the hypotheses of Lemma 3 are satisfied with $\Omega_1 = p$, $\Omega_2 = \psi$, and $f(w) = ag(w) + b[1 - g(w)]$. The Event C is then the event whose existence is asserted by Lemma 3.

Now it is possible to prove

Theorem 2: If there is an Event H for which $\psi(H) \neq p(H)$, then for any value of p within the limits $0 < P < 1$ there exist Events A and B satisfying the relations $p(A) = p$, $p(B) < p(A)$, $\psi(B) > \psi(A) + P(1 - P)|\psi(H) - p(H)|$.

Proof: It is sufficient to consider the case $p(H) \leqq 1/2$, $P \leqq 1/2$, $\psi(H) - p(H) = h > 0$; parallel arguments establish the result for other cases.

Suppose first that $0 < P \leqq p(H)$. By Lemma 4 (with $a = 0$ and $b = P/[1 - p(H)]$) there exists an Event A such that $p(A) = P$, while $\psi(A) = P[1 - \psi(H)]/[1 - p(H)]$. Similarly, taking $a = P/p(H)$, $b = 0$ yields the conclusion that there exists an Event B' with $p(B') = P$, $\psi(B') = P\psi(H)/p(H)$.

Subtraction yields the inequality:

$$\psi(B') - \psi(A) = \frac{hP}{p(H)[1 - p(H)]} > hP(1 - P)$$

Next suppose $p(H) < p \leqq 1/2$. The Event A may be selected as in the foregoing case, while an appropriate Event B' corresponds to $a = 1$, $b = [P - p(H)]/[1 - p(H)]$. Thus in this case also:

$$\psi(B') - \psi(A) = \psi(H) + \frac{P - p(H)}{1 - p(H)}[1 - \psi(H)] - P\frac{1 - \psi(H)}{1 - p(H)} = \frac{h}{1 - p(H)} > hP(1 - P)$$

Now let the values of a and b corresponding to the Event B' be multiplied by the factor:

$$k = 1/2\left[1 + \frac{hP(1 - P) + \psi(A)}{\psi(B')}\right]$$

and designate an associated event by B. Since $k < 1$, it follows that:

$$p(B) = kp(B') < P = p(A)$$

while

$$\psi(B) = k\psi(B') = 1/2[\psi(B') + hP(1 - P) + \psi(A)] > hP(1 - P) + \psi(A)$$

Thus the Events A and B have the desired properties and the proof is complete.

Theorem 2 could be strengthened, in that the expression $hP(1 - P)$ could be replaced by any continuous function $G(P)$ satisfying the inequalities:

$$G(P) < \frac{hP}{p(H)[1 - p(H)]}, \quad 0 < P \leqq \min[p(H), 1 - p(H)]$$

$$G(P) < \frac{h}{\max[p(H), 1 - p(H)]}$$

$$\min[p(H), 1 - p(H)] \leqq P \leqq \max[p(H), 1 - p(H)]$$

$$G(P) < \frac{h(1 - P)}{p(H)[1 - p(H)]}, \quad \max[p(H), 1 - p(H)] \leqq P < 1$$

The proof is changed only by the replacement of $hP(1 - P)$, where it appears, by $G(P)$.

Lemma 4 also makes it possible to draw a strong conclusion about the cardinality of the set of identifying characteristics.

Theorem 3: If there exists an Event H for which $\psi(H) \neq p(H)$, the set R of identifying characteristics is nondenumerably infinite.

Proof: Let the hypothesis of the theorem be satisfied for some H. Suppose for convenience that:

$$p(H) \leqq 1/2 \leqq 1 - p(H)$$

Then

$$ap(H) + b[1 - p(H)] = 1/2$$

if

$$b = [1/2 - ap(H)]/[1 - p(H)]$$

Since the foregoing relations imply that $0 \leqq b \leqq 1$ if $0 \leqq a \leqq 1$, Lemma 4 applies, and for each value of a in the indicated range there exists an Event C_a such that $p(C_a) = 1/2$, while:

$$\psi(C_a) = \frac{1/2[1 - \psi(H)] + a[\psi(H) - p(H)]}{1 - p(H)}$$

Thus distinct values of a are associated with distinct values of $\psi(C_a)$. Since $p(C_a)$ is the same for all a, the identifying characteristics $r(C_a)$ must likewise be distinct for distinct a's. Thus a subset of R can be put into one-to-one correspondence with the unit interval, and the theorem is established.

The general mathematical setup underlying Theorem 4 is explained in the body of the paper; it is not altogether the same as that underlying Theorems 2 and 3. Theorem 4 is concerned with bets of the form "with probability p you win \$$x$; with probability $1 - p$ no money changes hands." Specifically, it assumes the existence of indifference curves

in a plane of such bets, and inquires about the restrictions which a NASEU model puts on such curves.

Axioms 1 and 2 of this Appendix still apply, Axiom 3 is irrelevant, and Parts *a* and *b* of Axiom 4 are assumed (but *not* Part *c*). The following additional axioms are assumed.

Axiom 5: For the particular situations in which Theorems 4 and 5 are to be tested, $\psi(p)$ is a continuous, differentiable, monotonically increasing function of p.

Axiom 6: For the particular situations in which Theorems 4 and 5 are to be tested, $u(x)$ is a continuous, differentiable, monotonically increasing function of x.

Axiom 7: The model asserted by Equation 1 of the text applies. That is, in this situation Bet A is preferred to Bet B if and only if:

$$\psi(p_A)u(x_A) > \psi(p_B)u(x_B)$$

Now we perform the construction described in the paper and illustrated in Figure 1. If S_A is the slope of the indifference curve passing through Bet A, we can prove

Theorem 4: $S_A S_C = S_B S_D$.

First comes a proof that if the indifference curves can be transformed into a form fitting Equation 4 of the paper, then the first half of Theorem 4 must hold. If Equation 4 can be written for any indifference curve, there must exist two functions g and h such that the equation for each indifference curve can be written:

$$g(x)h(p) = k_c \qquad [A1]$$

Differentiating Equation 11 gives:

$$g'(x)h(p)dx + g(x)h'(p)dp = 0$$

$$\frac{dp}{dx} = -\frac{g'(x)}{g(x)} \cdot \frac{h(p)}{h'(p)} \qquad [A2]$$

$$\frac{dp}{dx} = \Phi(x)\theta(p) = S(x, p) \qquad [A3]$$

which simply says that the slope of the indifference curve at any point may be written as a product of a function of x along and a function of p alone. Now it is tautologous that:

$$\frac{\Phi(x_1)\theta(p_2)}{\Phi(x_1)\theta(p_1)} = \frac{\Phi(x_2)\theta(p_2)}{\Phi(x_2)\theta(p_1)} \qquad [A4]$$

Substitution from Equation A3 reduces Equation A4 to:

$$\frac{S(x_1, p_2)}{S(x_1, p_1)} = \frac{S(x_2, p_2)}{S(x_2, p_1)} \qquad [A5]$$

But by definition (x_1, p_2) defines Bet A, (x_2, p_2) defines Bet B, (x_2, p_1) defines Bet C, and (x_1, p_1) defines Bet D. So Equation A5 may be rewritten:

$$S_A S_C = S_B S_D \qquad [A6]$$

The proof so far has shown that if Equation 4 is correct then Equation A6 follows. In short, Equation A6 is necessary for the correctness of Equation 4, and so of Axiom 7. It remains to show that it is also sufficient. The statement of Theorem 4 given in the text of the paper specifies that Equation A6 is sufficient as well as necessary for the correctness of Equation 4 only if at least one indifference curve has a negative slope; this restriction is necessary to rule out transformations on p and x which are monotonically decreasing rather than monotonically increasing. To prove sufficiency given at least one indifference curve with a negative slope, it is necessary to exhibit transformations which actually fit Equation 4.

To start, choose any point (x_0, p_0), $x_0 > 0$, $0 < p_0 < 1$. Now let

$$u = \exp - \int_{x_0}^{x} S(x, p_0)dx \qquad [A7]$$

$$\psi = \exp S(x_0, p_0) \int_{p_0}^{p} \frac{1}{S(x_0, p)} dp \qquad [A8]$$

Now it is necessary to prove that these transformations fit Equation 4 if they fit Equation A6.

First, differentiate Equations A7 and A8:

$$d\psi = \left[\exp S(x_0, p_0) \int_{p_0}^{p} \frac{1}{S(x_0, p)}\, dp\right] \times \left[S(x_0, p_0) \frac{1}{S(x_0, p)}\, dp\right]$$

$$du = \left[\exp - \int_{x_0}^{x} S(x, p_0) dx\right] \times [-S(x, p_0)dx]$$

So

$$\frac{d\psi}{du} = -\frac{\psi}{u}\cdot\frac{S(x_0, p_0)}{S(x, p_0)S(x_0, p)}\frac{dp}{dx} \quad [A9]$$

Substituting into A9 from A6:

$$\frac{d\psi}{du} = -\frac{\psi}{u}\cdot\frac{S(x_0, p_0)}{S(x, p)S(x_0, p_0)}\frac{dp}{dx}$$

$$\frac{d\psi}{du} = -\frac{\psi}{u}\cdot\frac{1}{S(x, p)}\cdot\frac{dp}{dx}$$

But $\frac{dp}{dx} = S(x, p)$ by definition. Therefore, $\frac{d\psi}{du} = -\frac{\psi}{u}$ which is the differential equation for the family of rectangular hyperbolas.

Psychological Review
1962, Vol. 69, No. 2, 136–148

DECISION MAKING:

OBJECTIVE MEASURES OF SUBJECTIVE PROBABILITY AND UTILITY[1]

GORDON M. BECKER[2]

General Dynamics Corporation, Groton, Connecticut

This paper describes several methods for measuring subjective probability. Some of the methods require a priori measures of utility, others measure both utility and subjective probability simultaneously. The measures, derived from choice behavior, do not necessarily reflect the subject's (*S*'s) conscious opinions or feelings. The subjective probability estimates may be useful as measures of confidence, bias, prejudice, or expectancy.

The basis for the methods is a modification of Edwards (1955) Subjective Expected Utility (SEU) model. It is assumed that:[3]

1. A person confronted with a choice between two or more alternatives selects the alternative that appears, to him, to offer the highest average satisfaction (SEU). The SEU of an alternative, A, is defined as:

$$\mathrm{SEU}(A) = \sum_{i=1}^{I} \psi(A_i)u(A_i)$$

where:
- A_i is the ith possible outcome of course of action A; $i = 1, 2, \cdots, I$.
- $\psi(A_i)$ is S's subjective probability that Outcome A_i will occur.
- $u(A_i)$ is the satisfaction (utility) S expects to derive if Outcome A_i occurs when S takes course of action A.

2. The sum of the subjective probabilities of a set of mutually exclusive and exhaustive events equals unity,[4] i.e.:

$$\sum_{i=1}^{I} \psi(A_i) = 1$$

3. Utility and subjective probability do not interact.
4. The satisfaction (utility) S associates with an event is measurable.
5. The measures of S's utility reflect the utilities operating during the subjective probability measurement session, i.e., the utilities have not changed during the period intervening between the utility and the subjective probability measurement sessions.

In addition to the above, the assumptions inherent in the method used to measure S's utilities must be added. If the utility measures are based upon the assumption that utility differences of all magnitudes are discernable (Coombs & Komorita, 1958; Edwards, 1955; Luce & Raiffa, 1957; Savage, 1954; Siegel, 1956), i.e., that any difference in SEU leads to 100% preference for the alternative with the higher SEU, this assumption becomes part of the present model.

[1] The major ideas discussed in the first half of this paper were presented at the Thirteenth Annual Meeting of the Eastern Psychological Association, Atlantic City, April 3, 1959. This investigation was supported in part by Research Grant M-2348(RI) from the National Institute of Mental Health, United States Public Health Service, in part by the General Dynamics Corporation, and in part by Itek Corporation.

[2] Now located at TEMPO, General Electric Company.

[3] The author is indebted to Ward Edwards for his help in making these assumptions explicit, for encouraging the author to pursue these ideas, and for suggesting that the author consider the effects of various degrees of information about utility upon the determination of subjective probability.

[4] This is a modification of Edwards SEU model in that he did not require that the sum of subjective probabilities equal unity.

However, if the utility measures are based upon the assumption that preferences are stochastic when utility differences are small (Coombs, 1959; Davidson & Marschak, 1959; Davidson, Suppes, & Siegel, 1957; Luce, 1959; Suppes & Walsh, 1959), then the stochastic assumptions become part of the present model.

Figure 1 shows the SEUs of two alternatives Y_1 and Y_2 as functions of subjective probability. The figure illustrates a situation in which you, as S, must choose between going to the airport (Y_1) or going to the Green Room (Y_2). In this example you want to meet Bill, who will be either (B_1) at the airport, or (B_2) at the Green Room. The satisfactions (utilities) you expect to derive from the four possible outcomes of this situation are given by ϕ_{11}, ϕ_{12}, ϕ_{21}, and ϕ_{22}.

The model predicts that you will choose Y_1 (the airport) whenever your subjective probability that Bill is at the airport, $\psi(B_1)$, is greater than X_1. Note that X_1 is the subjective probability at which the two alternatives have equal SEU and intersect. (If a stochastic model is used to determine utilities, X_1 will be a range rather than a point.) Note also that the sum of the subjective probabilities, $\psi(B_1) + \psi(B_2)$, is unity, indicating that you are sure Bill will be either at the airport or the Green Room, not at both, and not somewhere else.

YOU CHOOSE Y	BILL CHOOSES B	
	AIRPORT (B1)	GREEN Rm (B2)
AIRPORT (Y1)	$\phi(11)$	$\phi(12)$
GREEN Rm (Y2)	$\phi(21)$	$\phi(22)$

WHERE: $\phi(22) > \phi(11) > \phi(21) > \phi(12)$

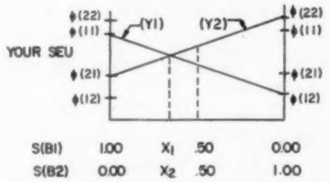


FIG. 1. Hypothetical problem illustrating SEU model.

MEASURING SUBJECT PROBABILITY

The subjective probability method proposed in this paper is based upon the reverse of the above interpretation, i.e., instead of predicting the alternative S chooses, the method estimates S's subjective probability at the time of choice. From Figure 1, it is evident that when S chooses Alternative Y_1, his subjective probability $\psi(B_1)$, must be between X_1 and unity, whereas when S chooses Y_2, $\psi(B_1)$ must be between X_1 and 0. Knowledge of both the utilities and the choice permits one to place upper and lower bounds on $\psi(B_1)$.

The range between the upper and lower bounds can be decreased by having S choose from among several alternatives. In general, if N different utilities are known, it is possible to construct N^2 different two-outcome alternatives.

Let $(a_i, p; a_j)$ represent an alternative, A, that pays a_i when p occurs and a_j otherwise. If $u(a_1) > u(a_2) > \cdots > u(a_n)$ it follows that $(a_1, p; a_1)$ is preferred to all other alternatives at all probabilities except $\psi(p) = 1$ at which point all alternatives of the form $(a_1, p; x)$ are equally preferred and except at the point $\psi(p) = 0$ when all alternatives of the form $(x, p; a_1)$ are equally preferred. Alternative $(a_1, p; a_2)$ is preferred to all alternatives of the form $(a_1, p; x)$ when $u(x) < u(a_2)$ and is preferred to all alternatives of the form $(x, p; a_2)$ when $u(x) < u(a_1)$. However, there are $N - 1$ alternatives of the form $(x, p; a_1)$ which are preferred to $(a_1, p; a_2)$ when $\psi(p)$ is small, but are not preferred when $\psi(p)$ is large.

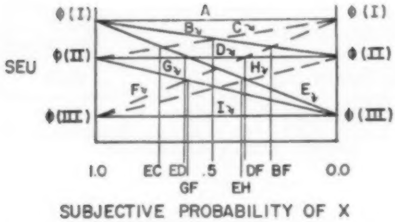


FIG. 2. Example illustrating relationship between SEU, subjective probability, and alternative choices.

Hence, these $N-1$ alternatives intersect $(a_1, p; a_2)$ at various $\psi(p)$. In general, alternatives of the form $(a_i, p; a_j)$ intersect $N-i$ alternatives of the form $(x, p; y)$, when $u(x) < u(a_i)$ and $u(y) > u(a_j)$. Figure 2 shows the nine two-outcome alternatives that can be formed from three distinct utilities and the nine subjective probability intersection points that result. In general, N distinct utilities will result in N^2 two-outcome alternatives and $N^2(N-1)^2/4$ intersections.

Let the subjective probability at the point at which Alternative A intersects Alternative B be $\psi(A, B)$—where $A=(a, p; a')$ and $B=(b, p; b')$. Since the SEUs of A and B are equal at the intersection point, it follows that:

$$\psi(p)u(a) + [1-\psi(p)]u(a') = \psi(p)u(b) + [1-\psi(p)]u(b')$$

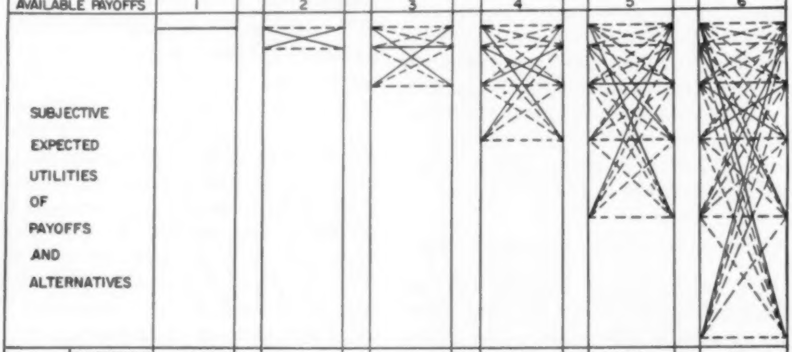
and:

$$\psi(A, B) = \frac{u(a')-u(b')}{u(b)-u(b')-u(a)+u(a')}$$

It is evident that $\psi(A, B) = \psi(B, A)$, i.e., that reversing the order of the alternatives does not change the subjective probability at the intersection point.

However, reversing the payoffs (utilities) within an alternative will change the crossover point. Let Alternative A' be $(a', p; a)$ and let Alternative B' be $(b', p; b)$. It follows that:

$$\psi(A', B') = \frac{u(a)-u(b)}{u(b')-u(b)-u(a')+u(a)}$$



AVAILABLE PAYOFFS		1	2	3	4	5	6
NUMBER OF PREFERENCE RANKS REQUIRED OF SUBJECT	ALL POSSIBLE	1	2	8	32	92	210
	FIVE	1	2	4	6	10	14
	FOUR	1	2	4	6	10	14
	THREE	1	2	4	6	10	12
	TWO	1	2	4	6	8	10
	ONE	1	2	3	4	5	6
		MAXIMUM NUMBER OF OBTAINABLE SUBJECTIVE PROBABILITY INTERVALS					

FIG. 3. The number of obtainable subjective probability intervals as a function of the number of required choices and the number of available payoffs.

Comparing $\psi(A', B')$ with $\psi(A, B)$ it will be seen that $\psi(A', B') = 1 - \psi(A, B)$.

It is also evident, from the above equations, that $\psi(A, A') = 1/2$, i.e., reversing the payoffs within an alternative will result in the reversed alternative intersecting the original at $\psi(p) = 1/2$. In general, $N(N - 1)/2$ of the intersections of the two-outcome alternatives formed from N unique utilities will intersect at $\psi(p) = 1/2$. One need know only the ranks of the two utilities used in these alternatives to determine which alternative has the higher SEU at subjective probabilities greater (or less) than 1/2.

The maximum number of unique intersections of two-outcome alternatives that can be achieved from N unique utilities is therefore:

$$\frac{(N^2 - 1)(N^2 - 2N) + 4}{4}$$

The maximum number of subjective probability intervals that can be determined from N unique utility points is one more than the number of unique intersections, as shown in Figure 3.

Utilities Measured on Interval or Ratio Scales

One method for estimating the subjective probability operating at the time of S's choice is to have S select one of several alternatives. For example, if S must choose between Alternatives E, D, and F shown in Figure 2 and he chooses Alternative E, his subjective probability, according to the model, must be between ED and unity. Whenever S selects Alternative F, his subjective probability must be between DF and 0; and when S selects Alternative D, his subjective probability must be between ED and DF. If the utilities, $\phi(\mathrm{I})$, $\phi(\mathrm{II})$, and $\phi(\mathrm{III})$ are measured on a ratio or interval scale, the numerical values of these probability bounds can be determined from the equations for the intersections of the alternatives.

The range between the upper and lower bounds can be reduced by having S again choose between alternatives. If S's original choice is E, Alternatives E and C might be presented. Given the initial preference E, if S now selects E over C his subjective probability must be between unity and EC, whereas a choice for C at this point would indicate a subjective probability between EC and ED. Similarly, Alternatives E and F might be presented if the original choice were D, and Alternatives B and F might be presented if the original choice were F. Given three utilities, it is possible, from two responses, to estimate which of six intervals contains S's subjective probability. A third choice would permit the experimenter to determine which of eight subjective probability ranges contains S's subjective probability.

Another technique for obtaining increased precision requires only a single response. S ranks his preferences for many alternatives instead of simply selecting the one he prefers most. Figure 3 shows the maximum number of intervals obtainable with from one to six utility points. Notice that the number of intervals depends upon the depth of the ranking required of S and the number of utility points used.

Utilities Measured on an Ordinal Scale

If only the ranks of the utilities are known, at least $2N - 2$ ranked probability intervals can be obtained. As indicated earlier, knowledge of the ranks of two utilities is sufficient information to enable the experimenter to construct a pair of alternatives such that S's choice indicates whether

his subjective probability is greater than, equal to, or less than 1/2. Alternatives B and C, and E and F, and G and H (Figure 2) are examples of such alternatives.

In order to obtain $2N - 2$ ranked probability from N different utilities when only the ranks of the utilities are known, construct a set of $2N - 3$ two-outcome alternatives. Let one alternative, AN, consist of winning the highest utility, a, when Event p occurs, and the lowest utility, n, otherwise, i.e., $AN = (a, p; n)$. From each of the other $N - 2$ utilities construct two alternatives. Let one alternative from each utility be a sure thing, and let the other alternative consist of winning the next lower utility when p does not occur, e.g., from Utility i construct Alternatives II and IJ where $II = (i, p; i)$ and $IJ = (i, p; j)$ and where the rank of $u(i)$ is one higher than the rank of $u(j)$. It is evident that Alternative II intersects AN at a higher subjective probability than the intersection of AN and IJ, i.e., $\psi(AN, II) > \psi(AN, IJ)$. Moreover, it is evident that $\psi(AN, BB) > \psi(AN, BC) > \psi(AN, CC) > \cdots > \psi(AN, MM)$ where $u(a) > u(b) > \cdots > u(m) > u(n)$. Two additional alternatives can be formed to complete the set of $2N - 2$ alternatives. One of these alternatives, $BA = (b, p; a)$, will intersect AN at the highest subjective probability, and the other alternative, $NM = (n, p; m)$ will intersect AN at the lowest subjective probability. Therefore, the $2N - 3$ alternatives formed in this manner intersect AN at $2N - 3$ different subjective probabilities. The ranks of the subjective probabilities at the intersection points are strictly determined by the ranks of the utilities. Hence, from N different utilities it is possible to construct a set of alternatives such that $2N - 2$ subjective probability intervals can be ranked from a knowledge of only the utility ranks.

A second technique for measuring subjective probability when only the utility ranks are known makes it possible to measure both utility and subjective probability simultaneously. Let the N utilities be ranked as before: $u(a) > u(b) > \cdots > u(n)$. Construct the following N two-outcome alternatives: $AN = (a, p; n)$; $BM = (b, p; m)$; $CL = (c, p; l)$; $\cdots$; $NA = (n, p; a)$. These N alternatives intersect, at most, at $(N^2-2N+2)/2$ distinct probability points to form $(N^2-2N+4)/2$ probability intervals. Consider the case of four utilities, ranked a highest and d lowest. Let the four two-outcome alternatives, constructed in accordance with the above plan, be: $A = (a, p; d)$; $B = (b, p; c)$; $C = (c, p; b)$; and $D = (d, p; a)$. The difference between Utilities a and b, i.e., $u(a) - u(b)$, must be less than, equal to, or greater than the difference between Utilities c and d. But when $u(a) - u(b) < u(c) - u(d)$, the ranks of the SEUs of the Alternatives A, B, C, and D are:

$$
\begin{array}{lll}
D>C>B>A & \text{when} & 0 \leq \psi(p) < \psi(C, D) \\
D=C>B>A & \text{when} & \psi(p) = \psi(C, D) \\
C>D>B>A & \text{when} & \psi(C, D) < \psi(p) < \psi(B, D) \\
C>D=B>A & \text{when} & \psi(p) = \psi(B, D) \\
C>B>D>A & \text{when} & \psi(B, D) < \psi(p) < 1/2 \\
C=B>D=A & \text{when} & \psi(p) = 1/2 \\
B>C>A>D & \text{when} & 1/2 < \psi(p) < \psi(A, C) \\
B>C=A>D & \text{when} & \psi(p) = \psi(A, C) \\
B>A>C>D & \text{when} & \psi(A, C) < \psi(p) < \psi(A, B) \\
A=B>C>D & \text{when} & \psi(p) = \psi(A, B) \\
A>B>C>D & \text{when} & \psi(A, B) < \psi(p) \leq 1
\end{array}
$$

Note that only 8 of the 24 possible rankings of A, B, C, D occur when $u(a) - u(b) < u(c) - u(d)$.

When $u(a) - u(b) = u(c) - u(d)$, the SEU ranks are:

$$
\begin{array}{lll}
D>C>B>A & \text{when} & 0 \leq \psi(p) < 1/2 \\
A=B=C=D & \text{when} & \psi(p) = 1/2 \\
A>B>C>D & \text{when} & 1/2 < \psi(p) \leq 1
\end{array}
$$

Note that all 24 rankings of A, B, C, D are possible when $u(a) - u(b) = u(c)$

$- u(d)$, but only two rankings occur when $\psi(p)$ is different from 1/2.

When $u(a) - u(b) > u(c) - u(d)$, the SEU ranks are:

$D > C > B > A$ when $0 \leq \psi(p) < \psi(A, B)$
$D > C > B = A$ when $\psi(p) = \psi(A, B)$
$D > C > A > B$ when $\psi(A, B) < \psi(p) < \psi(A, C)$
$D > C = A > B$ when $\psi(p) = \psi(A, C)$
$D > A > C > B$ when $\psi(A, C) < \psi(p) < 1/2$
$D = A > C = B$ when $\psi(p) = 1/2$
$A > D > B > C$ when $1/2 < \psi(p) < \psi(B, D)$
$A > B = D > C$ when $\psi(p) = \psi(B, D)$
$A > B > D > C$ when $\psi(B, D) < \psi(p) < \psi(C, D)$
$A > B > C = D$ when $\psi(p) = \psi(C, D)$
$A > B > C > D$ when $\psi(C, D) < \psi(p) \leq 1$

Note that only 8 of the possible 24 rankings of A, B, C, D occur when $u(a) - u(b) > u(c) - u(d)$.

From the above, it is evident that, when $\psi(p)$ is different from 1/2, six rankings of the alternatives occur only when $u(a) - u(b) > u(c) - u(d)$ and six other rankings occur only when $u(a) - u(b) < u(c) - u(d)$. Thus if S assigns any of these 12 distinctive rankings when $\psi(p)$ is not 1/2, the experimenter can immediately rank the subjective probability of p operating at that time and, at the same time, determine whether $u(a) - u(b)$ is greater or less than $u(c) - u(d)$. If $u(a) - u(b)$ is greater (less) than $u(c) - u(d)$, 10 of the 24 rankings will never occur, and if the utility differences are equal then 22 of the 24 rankings will never occur when $\psi(p)$ is different from 1/2. Therefore when subjective probabilities different from one-half are used, the experimenter can immediately determine from S rankings, whether $u(a) - u(b)$ is greater than or less than $u(c) - u(d)$, can immediately order the six subjective probability intervals obtainable from the four utilities, and can simultaneously determine which of the six intervals contains the $\psi(p)$ operative at the time of S's ranking.

Thus, given only the ranks of the utilities, it is possible to order the utility differences and the subjective probability intervals from S's rankings of the alternatives.

However, the above method does not allow for stochastic choice. If S ranks the alternatives consistent only with $u(a) - u(b) > u(c) - u(d)$ at one time, and consistent only with $u(a) - u(b) < u(c) - u(d)$ at another time, the method will lead to contradictory results. If people show such inconsistencies in their rankings it will be necessary to modify the method by extending it to include stochastic rankings of alternatives.

PRELIMINARY TEST OF THE GENERAL METHOD

Two preliminary experiments were conducted to explore whether or not Ss vary their selection of alternatives in accordance with the difficulty of a discrimination task and the objective probability of an event. A visual discrimination task was selected because it did not contain any explicit probability measures, because it provided a fairly critical test of the method, and because it permitted unambiguous control of the stimulus. Both experiments were conducted at Connecticut College[5] using the same 23 undergraduate women.

Each S was given a set of 45 problems from which her utility scale for points could be determined by linear programing (Davidson). A problem consisted of a choice between two alternatives. Each alternative consisted of winning 1500, 1000, 500, or 0 points or of losing 500 or 1000 points if Event ZEJ occurred; and winning or losing one of the above point scores if Event ZOJ occurred. The subjective probability of ZEJ (ZOJ) was assumed to be equal to 1/2. Linear programing solution could not

[5] The assistance of Mortimer H. Appelzweig and George O. Moeller in providing Connecticut College Laboratory facilities and students is gratefully acknowledged.

TABLE 1

DERIVED UTILITY SCALES

Subject	Objective Point Offers						ϕ
	−1000	−500	0[a]	500[a]	1000	1500	
QM	−1875	−1000	0	500	1375	1625	625
AG	−1061	−643	0	500	928	1142	71
MS	−2250	−1000	0	500	1000	1250	250
JN	−1500	−833	0	500	1000	1667	167
BW	−1333	−667	0	500	750	1083	83
GH	−1000	−750	0	500	750	1500	125
KT	−1333	−667	0	500	833	1333	167
CS	−1000	−500	0	500	875	1125	125
LB	−1250	−500	0	500	750	1250	250
EJ	−1300	−700	0	500	800	1200	100

[a] The zero points of all utility scales were arbitrarily set at zero objective points. The units of all utility scales were arbitrarily set by assigning 500 units to the distance between the utilities of zero and 500 objective points.

be found for many *S*s. The scales for 10 of the *S*s for whom solutions were found are shown in Table 1. Although the subjective distances between the point offers differed from the objective distances for most *S*s the rank-order of the points and the rank-order of the probabilities at which alternatives were preferred did not differ from the objective for the alternatives used in the two experiments.

Experiment I: Visual Discrimination

In the first experiment, *S* chose one of five alternatives on the basis of her ability to discriminate which of two line clusters contained the greater number of lines. Unknown to *S*, one of the two clusters always contained 100 lines. The other contained 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, or 125 lines. The lines were 1/8 to 1/4 inch long and were randomly distributed in an irregular, egg shaped pattern about 3/4 inch high and 1½ inches long. Both clusters were of the same irregular shape and area. Sixty pairs of these clusters were reproduced in mimeographed test booklets. Each booklet contained the same 10 pages with six problem pairs per page, but the order of the pages was randomized from one *S* to the next. The top cluster in each problem pair was labeled QUG and the bottom cluster was labeled QOB. Two of the QUG areas contained 75 lines, two contained 80, 85, 115, 120, and 125 lines. Four of the QUG areas contained 90, 95, 105, and 110 lines. Each of the 28 QUG areas was paired with a QOB area containing 100 lines. An additional 28 problems reversed the QOB and QUG areas. Four other problems were also presented in which QUG and QOB each contained 100 lines.

S was told that the test was designed to test her ability to estimate the number of objects in an area. The instructions informed *S* that she would be paid 1 cent for every 500 points she won, and that she would lose 1 cent for every 500 points she lost. However, *S* was guaranteed to net at least 1 dollar regardless of her losses. She could, if she selected properly, win up to 2 dollars. *S* won points by choosing one of the five

alternatives described below for each pair of QUG and QOB areas.

Alternative CIJ: You win 1500 points if there are more lines in area QUG than there are in area QOB, and you lose 1000 points if there are more lines in area QOB.

Alternative MEQ: You win 1000 points if there are more lines in area QUG than in QOB, and you lose 0 points if there are more lines in QOB.

Alternative ZEJ: You win 500 points for sure.

Alternative XIP: You win 1000 points if there are more lines in area QOB than in QUG, and you lose 0 points if there are more lines in QUG.

Alternative FEH: You win 1500 points if there are more lines in area QOB than in QUG, and you lose 1000 points if there are more lines in QUG.

These alternatives are shown in Table 2, together with the alternatives offered in the next two experiments.

Two lines were drawn to the right of each problem in the test booklet. One line was labeled Side A, and the other was labeled Side B. *S* wrote the name of the alternative she preferred on the Side A line. If she wished, she wrote the same name on the Side B line, or she wrote the name of a different alternative. *S* was told that only one of the alternatives indicated on the two lines would be considered in computing her score. This alternative was determined by flipping a coin labeled Side A on one side, and Side B on the other. If *S* wrote the same name on both lines, she was sure to have that alternative count towards her score. If she wrote different names on the two lines, one or the other would be selected, each with a 50/50 chance. This procedure distinguished between real preferences for an alternative and those forced by a choice between equally desirable alternatives. This method provides a seven-interval probability scale as shown in Table 3.

In order to avoid changing *S*'s utility for points during the experiment, *S* was not told which area contained more lines, which side of the coin appeared face up, or her point

TABLE 2

ALTERNATIVES OFFERED THE SUBJECT BY PROBLEM AND CONDITION

Problems	Condition	Alternatives				
	S chooses column alternative...........	CIJ	MEQ	ZEJ	XIP	FEH
Discrimination	If top area contains more lines, *S* receives	1500	1000	500	0	−1000
	If bottom area contains more lines, *S* receives	−1000	0	500	1000	1500
	S chooses column alternative...........	VUK	WUB	XEH	YOF	
First set	If ZEJ card is drawn, *S* receives	1500	1000	500	−1000	
Probability Deck	If ZOJ card is drawn, *S* receives	−1000	500	1000	1500	
	S chooses column alternative...........	QJF	FHJ	ZBJ	DJQ	
Second Set	If XIH card is drawn, *S* receives	−1000	0	1000	1500	
Probability Deck	If XEH card is drawn, *S* receives	1500	1000	0	−1000	

TABLE 3

PROBABILITIES ASSOCIATED WITH DISCRIMINATION RESPONSES

Response		Probability QUG contains more lines										
		Rational (Obj. Pts.)	Subjective (utility scale)									
One side	Other side		QM	AG	MS	JN	BW	GH	KT	CS	LB	EJ
FEH	FEH	0– .33	0– .12	0– .19	0– .10	0– .31	0– .29	0– .43	0– .27	0– .20	0– .29	0– .24
FEH	XIP	.33	.12	.19	.10	.31	.29	.43	.27	.20	.29	.24
XIP	XIP	.33– .50	.12– .50	.19– .50	.10– .50	.31– .50	.29– .50	.43– .50	.27– .50	.20– .50	.29– .50	.24– .50
FEH or XIP or ZEJ	CIJ or MEQ or ZEJ	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
MEQ	MEQ	.50– .67	.50– .88	.50– .81	.50– .90	.50– .69	.50– .71	.50– .57	.50– .73	.50– .80	.50– .71	.50– .76
MEQ	CIJ	.67	.88	.81	.90	.69	.71	.57	.73	.80	.71	.76
CIJ	CIJ	.67–1.00	.88–1.00	.81–1.00	.90–1.00	.69–1.00	.71–1.00	.57–1.00	73–1.00	.80–1.00	.71–1.00	.76–1.00

score until she had completed all three phases of the experiments. The utility scale, the discrimination and the probability deck problems, discussed in the next section, were all administered during one group session that lasted 1½–2 hours.

If the SEU model and the method for measuring subjective probability are appropriate, we expect *S* to choose a high probability alternative (CIJ–CIJ) whenever she is sure that QUG contains more lines, to choose FEH–FEH whenever she is sure QUG contains less lines, and to vary her choice in the seven intermediate steps shown in Table 3 as her confidence in her ability to discriminate varies. One also expects a high correlation between *S*'s confidence, as measured by the subjective probability associated with her choice, and the difficulty of the discrimination task. Since the ranks of the subjective probability levels do not differ from the objective probability levels, for the 10 *S*s for whom utility scales were obtained (Table 3), a high correlation would be expected between the ranks of the objective probabilities associated with alternatives selected by *S* and the difficulty of the discrimination task. The magnitude of the difference between the number of lines in the two areas provides an objective basis for ranking the difficulty of the discrimination. Thus, if a significant correlation between the ranks of the stimulus differences and the ranks of the probability intervals associated with the alternatives is found, it would support the notion of using the proposed method to measure subjective probability, confidence in decisions, and stimulus ambiguity or difficulty.

Results. Two Kendall coefficients of concordance were calculated for each *S*, one for four replications of the five problems involving 100, 105, and

110 lines in QUG and 100, 105, and 110 lines in QOB; and the other for the four replications of the five problems involving 90, 95, and 100 lines in QUG and QOB.

Spearman average rank correlations obtained from these coefficients are shown, for each *S*, in Table 4. The "best" estimates of each *S*'s choice, on each problem, obtained by averaging the *S*'s four choices on each problem, were used to compute a coefficient of concordance for the group as a whole. The average Spearman rank correlation obtained from the coefficient is labeled Total in Table 4. Note that 29 of the 46 correlations are significant beyond the .05 level and that the totals are significant beyond the .01 level.

TABLE 4

CORRELATIONS BETWEEN OBJECTIVE STIMULUS SCALE AND SUBJECTIVE PROBABILITY

Subject	Task			
	Discrimination (5 alternatives)		Probability (4 alternatives)	
	One area contains more than 100 (100–105–110)	One area contains less than 100 (90–95–100)	First Set	Second Set
1	.38	.64**	.78**	.79**
2	.24	.71**	.74**	.78**
3	.30	.66**	.81**	.88**
4	.84**	.55*	.83**	.89**
5	.46*	.08	.83**	.40**
6	.70**	.61**	.87**	.87**
7	−.03	−.20	.68**	.55**
8	.24	.59**	.84**	.80**
9	.39	.57*	.86**	.30*
10	.64**	.30	.82**	.79**
11	.73**	.99**	.89**	.89**
12	.67**	.71**	.85**	.78**
13	.77**	.75**	.84**	.75**
14	.68**	.59**	.91**	.86**
15	.15	.47*	.89**	.89**
16	.11	.39	.56**	.66**
17	.61**	.18	.89**	.61**
18	.55*	.75**	.86**	.86**
19	.10	.39	.88**	.82**
20	.51*	.67**	.85**	.78**
21	.46*	.89**	.84**	.82**
22	.83**	.53*	.87**	.90**
23	−.10	.53*	.90**	.75**
Total	.834**	.871**	.87**	.76**

* Significant at .05 level.
** Significant at .01 level.

EXPERIMENT II: OBJECTIVE PROBABILITY

The results of a second pair of tests, which employed explicit probabilities rather than line areas, are also shown in Table 4. *S* was told that one card would be drawn from a deck containing 100 cards. Each card in the deck was labeled either ZEJ or ZOJ in the first set, and either XIH or XEH in the second set. *S* was told exactly how many ZEJ (or XIH) cards were in a particular deck. In the first set, she selected one of the four "Probability Deck" alternatives shown in Table 2 and described below:

VUK: You win 1500 points if ZEJ is drawn, and you lose 1000 points if ZOJ is drawn.
WUB: You win 1000 points if ZEJ is drawn and you win 500 points if ZOJ is drawn.
XEH: You win 500 points if ZEJ is drawn and you win 1000 points if ZOJ is drawn.
YOF: You lose 1000 points if ZEJ is drawn and you win 1500 points if ZOJ is drawn.

The objective probabilities at which these alternatives intersect are shown in Table 5. In the second set, she selected one of the four alternatives, shown in Table 2 and described below:

QJF: You lose 1000 points if XIH is drawn and you win 1500 points if XEH is drawn.
FHJ: You win 0 points if XIH is drawn and you win 1000 points if XEH is drawn.
ZBJ: You win 1000 points if XIH is drawn and you win 0 points if XEH is drawn.
DJQ: You win 1500 points if XIH is drawn and you lose 1000 points if XEH is drawn.

The probabilities at which these alternatives intersect are shown in Table 5.

TABLE 5

THE PROBABILITIES AT WHICH ALTERNATIVES INTERSECT

Response		Rational (Obj. Pts.)	Subjective (utility scale)									
One side	Other side		QM	AG	MS	JN	BW	GH	KT	CS	LB	EJ
Probability ZEJ will be drawn												
YOF	YOF	0– .33	0– .10	0– .14	0– .08	0– .25	0– .20	0– .33	0– .21	0– .14	0– .22	0– .18
YOF	XEH	.33	.10	.14	.08	.25	.20	.33	.21	.14	.22	.18
XEH	XEH	.33– .50	.10– .50	.14– .50	.08– .50	.25– .50	.20– .50	.33– .50	.21– .50	.14– .50	.22– .50	.18– .50
YOF or XEH	VUK or WUB	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
WUB	WUB	.50– .67	.50– .90	.50– .86	.50– .92	.50– .75	.50– .80	.50– .67	.50– .79	.50– .86	.50– .78	.50– .82
WUB	VUK	.67	.90	.86	.92	.75	.80	.67	.79	.86	.78	.82
VUK	VUK	.67–1.00	.90–1.00	.86–1.00	.92–1.00	.75–1.00	.80–1.00	.67–1.00	.79–1.00	.86–1.00	.78–1.00	.82–1.00
Probability XIH will be drawn												
QJF	QJF	0– .33	0– .12	0– .19	0– .10	0– .31	0– .29	0– .43	0– .27	0– .20	0– .29	0– .24
QJF	FHJ	.33	.12	.19	.10	.31	.29	.43	.27	.20	.29	.24
FHJ	FHJ	.33– .50	.12– .50	.19– .50	.10– .50	.31– .50	.29– .50	.43– .50	.27– .50	.20– .50	.29– .50	.24– .50
QJF or FHJ	ZBJ or DJQ	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
ZBJ	ZBJ	.50– .67	.50– .88	.50– .81	.50– .90	.50– .69	.50– .71	.50– .57	.50– .73	.50– .80	.50– .71	.50– .76
ZBJ	DJQ	.67	.88	.81	.90	.69	.71	.57	.73	.80	.71	.76
DJQ	DJQ	.67–1.00	.88–1.00	.81–1.00	.90–1.00	.69–1.00	.71–1.00	.57–1.00	.73–1.00	.80–1.00	.71–1.00	.76–1.00

Each set of problems was constructed from 19 different decks. These decks contained the following numbers of ZEJ-ZOJ(XIH-XEH) cards: 0–100, 10–90, 15–85, 20–80, · · ·, 75–25, 80–20, 85–15, 90–10, 100–0. Thirty problems were constructed from each set of 19 decks by presenting 11 decks twice and 8 decks[6] once. Each set of 30 problems was included on a separate page in the test booklet after the discrimination experiment problems and the utility scale problems.

S again indicated the one, or two, alternatives she preferred, using the same Side A and Side B procedure. She was not informed of her point score on either set, nor of her point score on the discrimination or utility problems during the experiment.

Results. Kendall tau correlations between the objective probability ranks of the decks and the ranks of the alternatives selected by *S*, assuming *S* is rational, were computed for each *S* on each of the two problem sets. The results are shown in Table 4. The correlations labeled Total are average Spearman rank correlations obtained from Kendall's coefficient of concordance. The four entries in each row in Table 4 were obtained from the same *S* during the same experimental session.

Note that all 46 tau correlations computed for the probability problems are significant beyond the .05 level, and 45 of the 46 are significant beyond the .01 level. The average Spearman rank correlations are .87 for the first set and .76 for the second set. These results indicate that *S*s tended to select high subjective probability alternatives for highly likely events, and low subjective probability alternatives for events of low probability.

[6] The eight decks presented only once were the 100/0, 85/15, 80/20, 70/30, 30/70, 15/85, 10/90, and 0/100.

Summary of Experimental Results

The empirical results in both the discrimination and probability tasks support the use of the proposed method as a means for measuring subjective probability, confidence, and bias, and suggest that the method may provide a means for measuring stimulus ambiguity. The high correlations also indicate that it is an acceptable procedure for use in individual prediction. The method provides an estimate of subjective probability from a single response.

The fact that the correlations are less than 1.0, however, indicates that *S*s do not consistently choose the same response to the same stimulus. This may reflect fluctuations in utility during the experiment, or may result from obtaining subjective probability measurements that were more precise than those employed by *S*, or the results may be due to real changes in *S*'s expectancy from one stimulus presentation to the next. The fluctuations may further have been due, in this experiment, to the alternatives having approximately equal SEU over a wide range of discrimination problems and objective probabilities.

CONCLUSIONS

This paper described several methods for obtaining measures of subjective probability from a single response. One method required a priori knowledge of *S*'s utility scale and the other required that *S* never rank a set of alternatives inconsistently. Both methods were based on modifications of the SEU model and required that utility and subjective probability not interact.

Two experiments were conducted

to test one of the assumptions basic to both methods: the assumption that people vary their preference for alternatives according to subjective probability. The results supported this assumption, indicating that if utility and subjective probability do not interact, estimates of subjective probability can be obtained from a single response.

Several questions remain. The methods must be subjected to further empirical test. The reliability of the subjective probability estimates, the effects of increasing the number of alternatives presented to *S*, the relationship between subjective and objective probability, and the validity of the basic assumptions of the SEU model are still to be determined.

REFERENCES

COOMBS, C. H. Inconsistency of preferences as a measure of psychological distance. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theories.* New York: Wiley, 1959.

COOMBS, C. H., & KOMORITA, S. S. Measuring utility of money through decision-making. *Amer. J. Psychol.*, 1958, **71**, 383–389.

DAVIDSON, D., & MARSCHAK, J. Experimental tests of a stochastic decision theory. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theories.* New York: Wiley, 1959.

DAVIDSON, D., SUPPES, P., & SIEGEL, S. *Decision-making: An experimental approach.* Stanford: Stanford Univer. Press, 1957.

EDWARDS, W. The prediction of decision among bets. *J. exp. Psychol.*, 1955, **51**, 201–214.

LUCE, R. D. *Individual choice behavior.* New York: Wiley, 1959.

LUCE, R. D., & RAIFFA, H. *Games and decisions: An introduction and critical survey.* New York: Wiley, 1957.

SAVAGE, L. J. *The foundations of statistics.* New York: Wiley, 1954.

SIEGEL, S. A method for obtaining an ordered metric scale. *Psychometrika*, 1956, **21**, 207–216.

SUPPES, P., & WALSH, K. A nonlinear model for the experimental measurement of utility. *Behav. Sci.*, 1959, **4**, 204–211.

(Received August 23, 1960)

Psychological Review
1962, Vol. 69, No. 2, 149-155

THEORETICAL NOTES

A NOTE ON INTENSITY GENERALIZATION AND PROTHETIC SCALING [1]

SHELDON H. WHITE
The University of Chicago

Tests for stimulus generalization to variation in stimulus intensity have generally followed a typical procedure. A set of three, four, or five stimulus intensity values are selected. One group of subjects learns a response to the weakest intensity, another group learns to the strongest, and then both groups are tested for generalized responding to the complete set of stimuli.

The two-way gradients found by such a procedure have shown a typical asymmetry despite a respectable diversity of types of subject, stimulus dimensions, and learning paradigms. A description of the asymmetry given by Hull (1949) notes five points: (*a*) The curve of the group trained at the weak intensity value is concave downward; (*b*) The curve of the strong-trained group is concave upward; (*c*) The strong-trained group shows stronger response to its training stimulus than does the weak-trained group; (*d*) The downward curve is steeper in slope than the upward curve; (*e*) The upward and downward curves intersect at their midpoints.

Explanations of the asymmetry have involved an energizing property of CS intensity (Hull, 1949), or else a differential conditioning process involving contextual stimuli (Logan, 1954; Perkins, 1953). However, the shape which a curve will appear to have depends, of course, upon baseline and ordinate scaling. The purpose of the present note is to point out that: (*a*) newer psychophysical scales offer an alternative to the practice, traditional in intensity generalization studies, of selecting intensity values at equal logarithmic intervals; (*b*) arbitrary logarithmic scaling can produce generalization asymmetry resembling the typical picture under discussion; and (*c*) rescaling baselines of existing curves in terms of magnitude-estimation scales reduces the apparent asymmetry of some of the intensity generalization data.

Scaling of Prothetic Dimensions

Stevens (Stevens, 1957; Stevens & Galanter, 1957) has recently presented evidence for a dichotomy among stimulus dimensions. Those dimensions along which stimulus change appears to take place by addition or subtraction of stimulation are classed as prothetic, while other dimensions, for which change is to some extent a matter of substitution, are classed as metathetic. Prothetic dimensions are generally associated with intensity attributes. The stimulus dimensions along which asymmetry of generalization has been found (brightness, loudness, and area) are classified as prothetic, while those which have showed approximate symmetry (hue, pitch, locus on skin, locus in space, length of line) are either metathetic ***or have been treated as though they were.*** That is, experimental sets of stimulus values have not been logarithmically spaced in the manner usually employed for prothetic dimensions.

The logarithmic spacing of a series of intensity values arranges them so that approximately equal numbers of jnd's fall between successive values in the series, assuming that the stimuli so spaced are thereby also arranged in even steps

[1] Study performed with the aid of a Ford Foundation in Behavioral Sciences research grant. The writer wishes to express his gratitude to Donald W. Fiske, Austen H. Riesen, and Charles C. Spiker for helpful critical readings, and also to thank Paul Grim for computing the asymmetry indices of the two figures.

of similarity, or psychological distance. The plausibility of the assumption has led investigators of stimulus generalization to space their stimuli in accord with the principle; Hovland (1937a) actually counted jnd's to locate stimuli, and later studies have used the logarithmic spacing as a good approximation of this procedure.

Another psychophysical estimate of apparent stimulus distances produces an alternative scaling. The method of magnitude estimation requires the observer to assign numbers to stimuli in accord with what he judges their intensities to be. Differences between numerical magnitude estimations for stimulus intensities are typically in better agreement with the actual physical energy differences between those stimuli than the jnd curve might lead one to expect. That is, magnitude estimations are generally related to physical energy values by a power function, of less curvature than the logarithmic jnd function. Some dimensions even show a linear relationship between magnitude estimation and physical value (Stevens & Galanter, 1957).

While logarithmic (jnd) scaling has been traditional and accepted, magnitude scaling has some points in its favor. The summing of jnd's to estimate psychological distance rests on an unproven assumption, historically controversial, that jnd's are equal, summable, "units" of sensation. Magnitude scaling does not appear to require such a critical assumption. Further, magnitude scaling, and only this psychophysical technique to date, has showed some amount of validity by permitting quantitative cross-modality scaling predictions (Stevens, Mack, & Stevens, 1960).

EFFECTS OF LOGARITHMIC SCALING

What is the effect of baseline scaling on a two-way generalization curve? Suppose that five intensity values have been chosen so that the distances between their logarithms are equal. The line at the top of Figure 1 illustrates the placement of the stimuli, A-E, on the physical

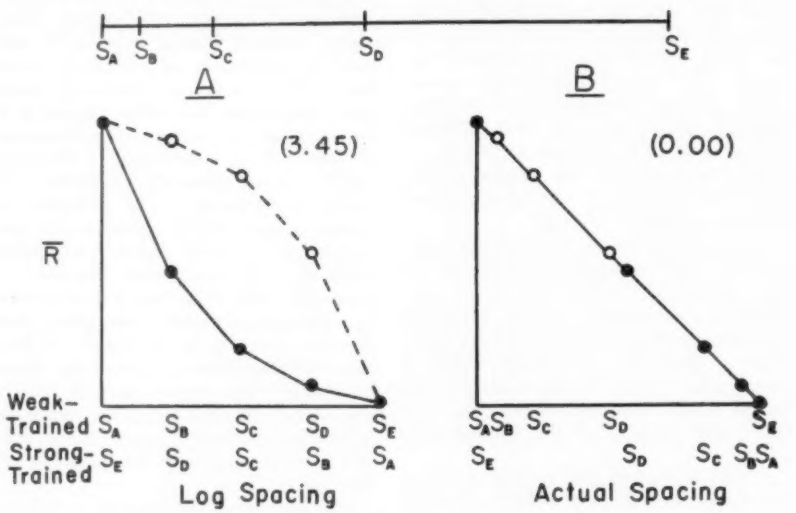


FIG. 1. Linear, symmetrical generalization curves inappropriately drawn over a logarithmically scaled baseline (Graph A), and drawn over the correctly scaled baseline (Graph B). (The training points for the weak-trained group—open circles, dashed line, and the strong-trained group—filled circles, solid line, are at the left, and the curves extend rightward. The parenthesized numbers are a rough index of asymmetry, representing the number of square inches between the curves as originally drawn in 3" × 3" size.)

dimension; the traditional assumption would be that they mark off approximately equal apparent steps.

A weak-trained group of subjects learns a response to S_A, while a strong-trained group learns to S_E, following which both groups are tested for generalized responding to the complete set of five stimuli. We will suppose that the generalization function is linearly, symmetrically related to physical energy differences between training and test stimulus; the subject generalizes in direct proportion to the physical energy differences between stimuli.

Hypothetical results under the above conditions are given in Figure 1, in both the A and B graphs. In the A graph, the spacing of the stimuli along the abscissa is even, following the assumption that equal logarithmic distances represent equal apparent stimulus change. The generalization curves appear to be markedly asymmetrical, with that for the weak-trained group being concave downward and that for the strong-trained group being concave upward.

In the B graph, the stimuli are spaced according to their physical energy distances, and the curves for the weak-trained and strong-trained groups are symmetrical and linear.

It is possible that the practice of logarithmic spacing, characteristic of intensity generalization studies, may have analogously led to an observed asymmetry. Some support for this argument will now be sought in the existing literature.

The analysis of the hypothetical example just given is to serve as a model for what follows. In that example, it was assumed that apparent stimulus differences were in exact accord with the physical energy differences between stimuli. For the studies to be considered, it will be assumed that apparent stimulus differences are best estimated from the scales yielded by the psychophysical method of magnitude estimation.

STUDIES OF INTENSITY GENERALIZATION

The general format of a study of bidirectional intensity generalization has been described above. Some distinction should be made between a "pure" design—in which each subject is tested on only one stimulus for generalization—and the more common counterbalanced design, where each subject is tested on all of the stimuli, with different subgroups of subjects used to counterbalance the order of testing among stimuli. Despite the fact that the data from counterbalanced designs are more difficult to interpret—being a mixture of generalization, extinction, retraining, and generalization of extinction effects—such designs have been used because of the considerable economy in number of subjects which they permit.

The relevant literature may now be briefly described:

1. Hovland (1937a) used a GSR conditioning procedure, with his stimuli being four 1000 cycle tones separated in loudness by 50 jnd steps. Separate groups of 16 college students were trained at his faintest and loudest tones, and then tested by a counterbalanced procedure.

2. Brown (1942) used rats in a straight alley approach situation. His stimuli were three illuminations of a large ground-glass screen at the end of the alley. There were two groups of 18 rats each, tested by a counterbalanced procedure.

3. Grice and Saltz (1950) also used rats in a straight alley approach situation. Their stimuli were five white circles, four of them graded in area at logarithmically equal intervals and the fifth at an intervening half-step. Testing was by a pure procedure, with 15 subjects for each test stimulus. Two sets of data are to be considered here, the resistance to extinction data reported originally, and speed data reported later (Grice, 1956).

4. Miller and Greene (1954) used rats in a straight runway avoidance paradigm. The stimuli were three loudnesses of a buzzer. There were 30 rats, tested by a pure procedure. Each data point of the resistance to extinction measure was based upon five animals.

5. Spiker (1956) tested children on a task in which they responded freely dur-

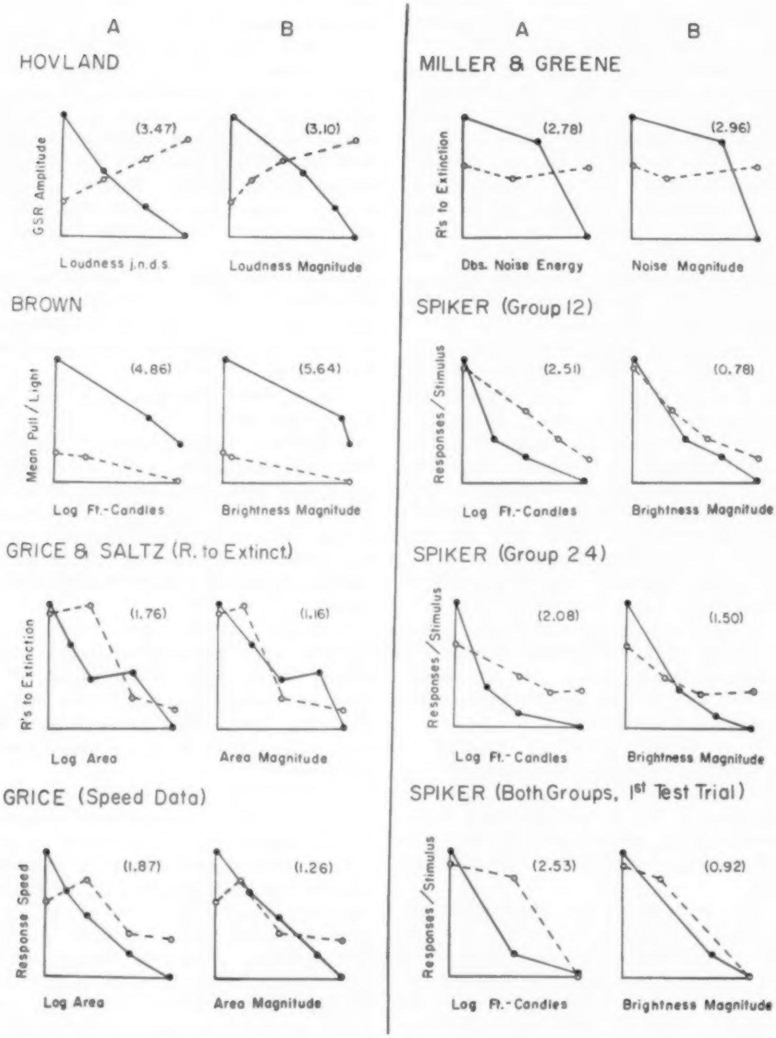


FIG. 2. The A columns show the results of intensity generalization studies with abscissae scaled to log stimulus energy (or, in Hovland's case, jnd distances). The training points for the weak-trained groups (open circles, dashed lines), and the strong-trained groups (filled circles, solid lines) are at the left, and all curves extend rightward. In the B columns, the same data are presented over abscissae rescaled in terms of psychophysical magnitude estimations. (All figures have been proportioned to occupy the same square area. The parenthesized numbers are a rough index of asymmetry, representing the number of square inches between the curves as originally drawn in 3″ × 3″ size.)

ing 3-second trial intervals. His stimuli were four illuminations of a circular ground-glass stimulus window; three of the stimuli were at logarithmically equal foot-candle intervals and the fourth at an intervening half-step. Thirty children were given 12 trials of training (15 weak-trained and 15 strong-trained), followed by a counterbalanced testing procedure; another 30 were given 24 trials of training before test. These two training groups will be considered separately below. In addition, the data of the pooled groups on their first generalization test trial will be considered, as a relatively pure generalization measure.

An additional study (Heyman, 1957) is clearly relevant in design but will not be further considered because the data, as tabled, did not permit clear appraisal of curve shapes.

The A columns of Figure 2 show the generalization curves reported for the above studies. To provide a clearer picture of the degree of symmetry or asymmetry of the curves, the training points of both the weak-trained and strong-trained groups have been placed on the left, and the curves extended rightward. All figures have been proportioned to occupy the same square area. The scaling along the abscissae of the Column A figures is logarithmic, except in the case of Hovland, where it is in terms of jnd units.

The ordinates of the Column B figures are identical with those of their Column A counterparts; the abscissae of the Column B figures, however, have been rescaled in the terms of psychophysical magnitude estimations. The stimulus energy values (E) for each study were transformed to magnitude scale values (M) using the power function $M = E^k$, where the exponent k is a characteristic of the dimension established by psychophysical investigation. For Hovland's loudnesses of 1000 cycle tones, k was 0.3 (Stevens, 1956). For Brown's study, as for Spiker's, brightness of illumination called for a k of 0.36 (Stevens & Galanter, 1957). Miller and Greene's buzzer loudnesses were converted using an equation for white noise (Stevens, 1958) with a k of 0.3. Finally, Grice's dimension of visual area involved a k range of 0.9–1.15 (Stevens, 1957); a k of 1.0 was used here.

If the asymmetry of the Column A figures were the result of baseline scaling alone, their treatment in Column B should make them roughly symmetrical. Several extraneous factors in the data being considered combine to make this a fairly gross specification.

First, the psychophysical data on which the rescalings were based comes from adult human observers, but three of the five studies were done with rats and a fourth with children. The Column B baselines are therefore only educated guesses at a proper scale for those subjects.

Second, the response strengths of the weak-trained and strong-trained groups were not always equal at their respective training stimuli. Generalization curves extending from training stimuli of different height would not be superimposed, nor is there any good reason to expect them to be parallel. Connected in here is an ordinate scaling problem—the probability that the same amount of response change has different implications starting from different levels of responding.

A third factor which might prevent rescaling from bringing about full symmetry applies particularly to counterbalanced designs. If the line at the top of Figure 1 is examined, it will be seen that logarithmic spacing of stimuli causes those of weaker intensity to be placed relatively closer together than stimuli of greater intensity. During repeated testing, the net effect of this is to produce less generalization of inhibition at the training value for the strong-trained group than for the weak-trained group, while loading relatively more generalized inhibition onto the tail of the strong-trained group. The strong-trained group consequently shows a steeper gradient than does the weak-trained group. This difference in steepness is not removed by rescaling. It might help account for differences between bright-trained and dim-trained subjects at their respective train-

ing values, the factor discussed in the preceding paragraph.

Now, examining Figure 2B, it is seen that the rescaling meets with partial success in symmetrizing the curves. Five of the curves—the two derived from the Grice data, and the three derived from the Spiker study—appear to be brought closer to symmetry. These five curves represent only three completely independent sets of data among them. On the other hand, the rescaling of the Hovland, Brown, and Miller and Greene studies has worked no improvement.

Discussion of Studies

Despite the fact that Hull advanced his five rubrics from knowledge of only the Brown and Hovland data, the failure to symmetrize either these or the Miller and Greene data is, to some degree, tied to discrepancies between these studies and Hull's typical picture. Hovland's weak-trained group shows a generalization curve regularly ascending away from the training point; Brown's weak-trained and strong-trained curves show no overlap; the Miller curves are exactly opposite in curvature to the standard. Only in the case of the Miller and Greene data—for which *N* is relatively low and reported response variance quite high—is there a clear cause to doubt the reliability of the reported curve shapes. Indeed, the ascending trend of Hovland's weak-trained group was reproduced by him in two later sets of intensity GSR generalization data (Hovland, 1937b).

On the other hand, the Grice and Saltz, and Spiker data, which on the whole provide the best complete approximation to Hull's picture, are most altered by rescaling. Their residual asymmetry might be due to the extraneous factors noted above, to sampling error, or to an intensity effect which is only reduced in apparent magnitude by rescaling.

Differentiae of studies not made symmetrical by rescaling might include: (*a*) a large difference between response strengths of the dim-trained and bright-trained groups to their training stimuli; (*b*) aversive, as opposed to appetitional, incentives; (*c*) auditory, as opposed to visual, stimuli; (*d*) counterbalanced as opposed to pure, designs; or (*e*) a stimulus set of greater "impact," whether because of intensity or of dominance as figure against ground.

What can be said, then, about the rescaling venture? There is no a priori reason for preferring one expression or picturing of a set of data to another, but choices can be made in terms of usefulness for interpretation and communication. Here, the limited reduction of asymmetry by magnitude scaling may usefully clarify where and how much a special intensity factor must be invoked. There is still evidence for an asymmetry of generalization, which may reflect an intensity factor. However, the intensity interpretation still faces as a difficulty the fact that a majority of the half-dozen conditioning studies which have searched for an effect of CS intensity have been unable to find it.

Summary

The typical asymmetry found in studies of two-way intensity generalization may involve baseline scaling, since all such studies have spaced their stimuli at equal jnd distances, either directly or by a logarithmic approximation. Magnitude estimations of stimuli yield a different but justifiable power-function spacing. An artificial example of a typical asymmetry produced by logarithmic spacing is given. Rescaling of intensity generalization data in the literature to magnitude-scale baselines results in increased symmetry in some cases.

REFERENCES

Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.

Grice, G. R. Response speed and size generalization. *Psychol. Rep.*, 1956, **2**, 246.

Grice, G. R., & Saltz, E. The generalization of an instrumental response to stimuli varying in the size dimension. *J. exp. Psychol.*, 1950, **40**, 702–709.

Heyman, W. Certain relationships between stimulus intensity and stimulus generalization. *J. exp. Psychol.*, 1957, **53**, 239–248.

Hovland, C. I. The generalization of conditioned responses: II. The sensory gen-

eralization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279–291. (a)

HOVLAND, C. I. The generalization of conditioned responses: III. Extinction, spontaneous recovery, and disinhibition of conditioned and of generalized responses. *J. exp. Psychol.*, 1937, **21**, 47–62. (b)

HULL, C. L. Stimulus intensity dynamism (*V*) and stimulus generalization. *Psychol. Rev.*, 1949, **56**, 67–76.

LOGAN, F. A. A note on stimulus intensity dynamism (*V*). *Psychol. Rev.*, 1954, **61**, 77–80.

MILLER, W. C., & GREENE, J. E. Generalization of an avoidance response to varying intensities of sound. *J. comp. physiol. Psychol.*, 1954, **47**, 136–139.

PERKINS, C. C. The relation between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, **46**, 225–231.

SPIKER, C. C. The stimulus generalization gradient as a function of the intensity of stimulus lights. *Child Develpm.*, 1956, **27**, 85–98.

STEVENS, J. C. Stimulus spacing and the judgment of loudness. *J. exp. Psychol.*, 1958, **56**, 246–250.

STEVENS, J. C., MACK, J. D., & STEVENS, S. S. Growth of sensation on seven continua as measured by force of handgrip. *J. exp. Psychol.*, 1960, **59**, 60–67.

STEVENS, S. S. The direct estimation of sensory magnitudes—loudness. *Amer. J. Psychol.*, 1956, **69**, 1–25.

STEVENS, S. S. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153–181.

STEVENS, S. S., & GALANTER, E. H. Ratio scales and category scales for a dozen perceptual continua. *J. exp. Psychol.*, 1957, **54**, 377–411.

(Received December 19, 1960)

Psychological Review
1962, Vol. 69, No. 2, 156-160

ALL-OR-NONE VERSUS INCREMENTAL LEARNING[1]

JOAN E. JONES

Defence Research Medical Laboratories; Toronto, Canada

The process of learning may be conceived of as the gradual strengthening of connections between events or, alternatively, as the establishment of such connections in an all-or-none fashion. Estes (1960) presented results which suggest that evidence for the all-or-none position tends to be obscured by the traditional techniques of grouping data. In one of a number of miniature experiments cited by Estes (1960, pp. 213–215) paired associates were recalled on successive test trials (T_1, T_2) after a single reinforced presentation (R_1). The sequence was R_1, T_1, rest, T_2. Estes argued that, if an incremental theory is tenable, the probability of a correct response appearing on T_2 should be independent of recall on T_1. An all-or-none theory suggests that correct responses on T_2 should occur only by chance unless a correct response was given on T_1 (see Figure 1). A close relationship was found between recall on T_1 and T_2, i.e., of the 49% of cases correct on T_1 75% were correct on T_2 (C-C cases) and of the 51% of cases incorrect on T_1 91% were incorrect on T_2 (N-N cases).

Estes' statement of incremental theory forces upon it the assumptions that a reinforced presentation results in equal increments in associative strength for all items in all subjects and that the strength of competing responses is equal in all cases. His deductions concerning the probability of a correct response on T_2 do not necessarily follow if these assumptions are not made. Estes contends that he has disposed of the point of item difficulty and individual differences by demonstrating that for his subjects and with his items a second reinforcement (R_2) interpolated between T_1 and T_2 yields approximately the same percentage of newly correct responses on T_2 (46%) as occurred on T_1 following R_1 (40%) (Estes, 1960, pp. 215–216). However, any prediction of the relative percentages of newly correct responses following one reinforced presentation and following two, involves making assumptions about the distributions of item difficulty and learning ability. The character of those distributions for such a limited number of items and subjects can be known only empirically. Although the approximate equivalence of the obtained percentages is consistent with an all-or-none interpretation, the demonstration does not at all preclude the possibility that R_1 results in unequal increments for different cases.

According to the latter view, the associations have achieved unequal strengths after R_1, hence they do not have equal probabilities of appearance. Subclassifying the cases into those which do (C) and those which do not (N) appear on T_1 biases the probabilities of appearance on T_2 without further reinforcement in favor of the C cases. Assuming that there is no forgetting between T_1 and T_2, the C-N and N-C cases represent approximately half of those in which the associations are at about threshold strength and the number of cases in each should be approximately equal. (The proportions will be equal in the special case where the C/N split on T_1 is 50/50.) Forgetting is reflected in an increase in the number of C-N cases and a decrease in the number of N-C cases. That is, incremental theory does not have to predict the frequency of N-C cases. However, it can predict trends in the N-C and C-N categories when the conditions of the experiment produce greater or lesser

[1] Defence Research Medical Laboratories Project No. 241, Report No. 241-4, PCC No. D77-94-01-16, H.R. No. 204.

degrees of forgetting. According to the all-or-none theory, forgetting will be reflected in an increase of cases in the C-N category but the frequency of cases in the N-C category will be unaffected unless subjects are guessing intelligently, in which case it will decrease.

Another problem which precludes uncritical acceptance of Estes' interpretation is the indeterminate theoretical chance level. According to an all-or-none theory the number of N-C cases should be no greater than that expected on the basis of sheer guessing. The chance value depends on the population of responses among which the subject is guessing, which might be a number less than, equal to or greater than eight, the number of items presented, and cannot be specified other than empirically. The alternative forms of guessing might be referred to as "intelligent," "unintelligent," and "stupid" guessing. Further information is necessary in order to evaluate whether the N-C cases which do occur, whatever their number, are sheer guesses or possible true recalls. This information is not available in the data as collected by Estes but an extension of his procedure may clarify why N-C cases do occur and also indicate whether associations have been established by R_1, though not revealed on T_1.

Three experiments are reported here. The first is a replication of Estes' experiment with the trial sequence R_1, T_1, "rest," T_2, confirming his result. The second is a variation and extension of Estes' procedure, the trial sequence being R_1, T_1, T_2, T_3, T_4. The third is a repetition of Experiment II using a double-blind technique.

EXPERIMENT I

Twenty-seven housewives acted as subjects in Experiment I. Each of eight associate-pairs, a consonant as stimulus and a number as response, were shown on separate 3″ × 2½″ cards. The responses were the numbers 1 through 8. The experimenter sat at a table opposite the subject and presented the cards manually after shuffling. On R_1

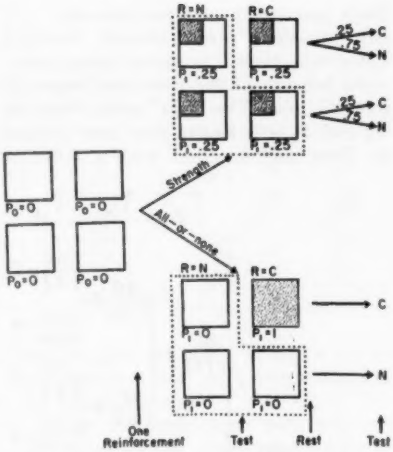


FIG. 1. Schema representing effects of a single reinforcement according to incremental (upper branch) vs. all-or-none (lower branch) theories. (Squares represent subjects with the proportion of darkened area in each indicating the probability of the correct response, C, for the given individual) (from Estes, 1960, p. 214.)

the cards were each held while the subject read aloud the consonant and the number. On test trials cards bearing only the consonant stimulus were shuffled and then presented. The test cards were held until the subject responded with the number she though correct or with a guess. No omissions were allowed. T_2 followed T_1 after approximately 7 minutes during which time the subject took part in another experiment which involved sorting cards bearing lines of different lengths and orientations. There was no warning that a second test would be made.

Experiment I was not an exact replication of Estes'. Apart from the usual differences of experimenter, subject, time, and place the present experiment differed from Estes' in four respects: (*a*) In the present experiment consonants rather than three-consonant syllables were used as stimuli. (*b*) Because the cards were shown manually the presentation and between-presentation duration were variable whereas

Estes timed both of these intervals. (*c*) No omissions were allowed whereas Estes was forced to accept some omissions because of the restricted intertrial interval. (*d*) The "rest" period between T_1 and T_2 was longer than that allowed by Estes and occupied with a different activity. From Estes' full report of his experiment (Estes, Hopkins, & Crothers, 1960) it appears that the "rest" interval was brief (estimated at 1½ minutes) and was used for a reinforced presentation of other items from the list. The probable unimportance of these procedural differences can be assessed from the results of Experiment I which are essentially the same as those obtained by Estes. Of the 53% of cases correct on T_1 75% were correct on T_2 (C-C) and of the 47% not correct on T_1 81% were not correct on T_2 (N-N).

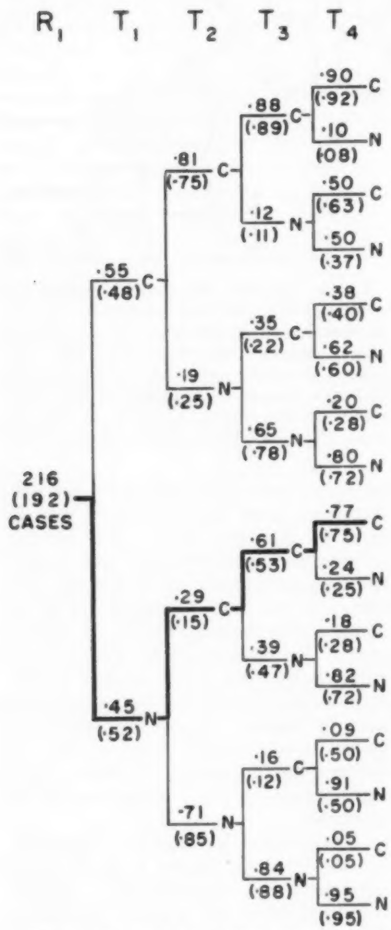


FIG. 2. Results of Experiment II and Experiment III (in parentheses). (Empirical values are proportions of instances in which correct, C, and incorrect, N, responses on a given test trial were followed by C and N responses on the following test trial.)

EXPERIMENT II

The subjects for Experiment II were the same 27 women who served in Experiment I. Both experiments were conducted in the same session, Experiment II first, and were separated by approximately 7 minutes during which time the subject took part in the card-sorting experiment. The procedure was as described for Experiment I except that the test cards were reshuffled and represented immediately until four test trials were completed. The experimenter gave no indication beforehand that more than one test would be made. The stimuli were consonants, none of which appeared in Experiment I. The responses were again the numbers 1 through 8.

The probability of recall on $T_{2\text{-}4}$ largely depends on whether the correct response is given on T_1 (see Figure 2), again confirming Estes' findings. However, features of the results are difficult to reconcile with an all-or-none theory.

Although the percentage of N-C cases (29%) is well within the range of chance variation, it is necessary to explain the tendency to repeat those guesses on T_3 and T_4, i.e., N-C-C and N-C-C-C cases. One would expect all-or-none theory to require the N-C-C and N-C-C-C values to approximate .29, i.e., the guessing rate established on T_2. The incremental interpretation demands that those values approximate .50, these cases being at threshold strength. The obtained values of .61

TABLE 1

TRENDS IN C–N AND N–C CASES AS A FUNCTION OF FORGETTING AND A COMPARISON OF N–C VALUES WITH THOSE EXPECTED ON THE BASIS OF INTELLIGENT GUESSING (g)

Experiment	Percentage correct		Retention index T_2/T_1	C-N	N-C	g
	T_1	T_2				
III	.55	.57	1.04	.19	.29	.22
II	.53	.48	.91	.25	.19	.21
Estes	.49	.39	.80	.29	.09	.20

and .77 are too high for either of these interpretations. These cases cannot be fully accounted for as part of a general tendency to repeat earlier calls on later trials. Of those who change their call on T_2, having been wrong on T_1, 61% of those correct repeat the second call once and 36% twice while of those again incorrect only 37% repeat the second call once and 14% twice. If the N-C cases are regarded as guesses it is necessary to explain why those calls are repeated with a much higher probability than incorrect responses. If they are regarded as true recalls the expected N-C-C and N-C-C-C values are .50, or less (if forgetting occurs), but in fact the values increase.

The implication is that there is some kind of feedback which distinguishes past responses as correct or incorrect. One possible source of such information under the conditions of Experiment II is the experimenter, who knows whether the response is correct or not. Experiment III was conducted to check this possibility. A new group of 24 subjects drawn from the same population learned the items used in Experiment II. One experimenter presented the paired-associates on R_1 and a second experimenter, who was unaware of the correct associations, presented the test cards on T_1 to T_4. The results of Experiment III appear in brackets in Figure 2. Again a high proportion of calls correct for the first time on T_2 were repeated on T_3 and T_4—a proportion considerably in excess of repetitions of incorrect calls. Of those correct on T_2 53% repeat the response once and 40% twice, while of those incorrect only 32% repeat the call once and 12% twice.

If the feedback does not originate from an unwitting experimenter then it must be somehow induced by the response itself. Bearing in mind that the associations in question are assumed to be at threshold strength and also the fact that recognition is easier than recall, it seems not unreasonable to suggest that in some cases the subject may recognize that the response she has just given as a guess was actually correct. Such an interpretation renders quite conceivable the notion that learning can occur on test trials. Estes et al. (1960) in fact noted such a phenomenon. However, the interpretation suggested does require that an association of liminal strength be set up on R_1 and is therefore inconsistent with an all-or-none interpretation.

Further, comparison of the results of Estes' experiment with those of Experiments II and III shows an effect which is also handled more readily by an incremental theory. If the total number of correct responses is taken as an index of retention then it is apparent that forgetting between T_1 and T_2 is greatest in Estes' experiment and least in Experiment III (see Table 1). The trend is consistent with the fact that Estes' subjects learned other associations between trials while the subjects in Experiment II sorted cards and in Experiment III there was no interpolated activity. From Table 1 it is also seen that as forgetting increases so the proportion of cases in

the C-N category increases, as predicted by both theoretical approaches, but also that the proportion in the N-C category decreases. A χ^2 test based on the frequencies in the N-C categories showed the latter effect to be significant at the .001 level of confidence. The effect cannot be accounted for in terms of an increase in the population of responses among which the subject is guessing. The expected N-C values for intelligent guessing were calculated as follows. If C is the proportion correct on T_1, c the probability that an item will be learned on R_1, and g the probability that the correct response will be guessed, then the equation for C (Estes et al., 1960) is:

$$C = c + (1 - c)g \qquad [1]$$

Substituting the obtained value of C and an arbitrary value of g (with the limitation that $g < C$) into Equation 1, solve for c. A new estimate of g is obtained from:

$$g = \frac{1}{n(1 - c)}$$

where n is the number of items in the list. The new estimate of g is then substituted into Equation 1 and the procedure reiterated until a stable value of g is obtained. The g values obtained after three iterations are shown in Table 1. A χ^2 test of the obtained N-C values against the frequencies predicted for intelligent guessing yielded a p of .001. That is, the drop in the N-C frequencies as a function of forgetting exceeds that expected simply as a decline in the chance level.

The data of Experiment II strongly suggest that subclassifying the cases according to their history on successive trials does no more than sort them in order of difficulty, each case being one item for one subject. The probability of appearance of a correct response on T_4 is: for cases formerly correct on all three trials (C-C-C cases), .90; for cases formerly correct on two of three occasions (C-C-N, C-N-C, and N-C-C), .50, .38, and .77 respectively; correct on one of three occasions (C-N-N-, N-C-N, and N-N-C), .20, .18, and .09 respectively; for those never before correct (N-N-N), .05. A similar ordering can be observed in the data of Experiment III. According to the all-or-none position the T_4 probabilities should be the same for all categories with the exception of C-C-C.

The writer appreciates Estes' statement that he does not mean to offer the experiment he reports as crucial tests of incremental versus all-or-none theory and agrees with his thesis that theorists should examine closely the extent to which their concepts are artifacts of experimental procedures. But insofar as the minature experiments are used as demonstrations, they are taken to create difficulties for incremental theory. Considered in the context of additional information from the experiments reported in this paper these same data raise difficulties for all-or-none theory. The criticism directed at the miniature experiments is not that they are too simple but that they are incomplete.

REFERENCES

ESTES, W. K. Learning theory and the new "mental chemistry." *Psychol. Rev.*, 1960, **67**, 207-223.

ESTES, W. K., HOPKINS, B. L., & CROTHERS, E. J. All-or-none and conservation effects in the learning and retention of paired associates. *J. exp. Psychol.*, 1960, **60**, 329-339.

(Received February 2, 1961)